Design and Synthesis of Conformational Switches and β-Turn Mimics Through Chemical Ligation

Selma Crecente García

Degree in Chemistry

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

School of Chemistry

College of Science and Engineering

University of Glasgow

August 2021

Abstract

The control of peptide conformation is crucial to facilitate further understanding of peptide and protein folding and activity. This thesis focuses on the control and mimicry of peptide secondary structure through two strategies: the design of conformational switching systems and the mimicry of β-turns through chemical ligation.

Chapter 1 describes the forces that determine protein structure and the conformations adopted by these biological systems. Strategies towards the design of β-hairpins, including β-turn mimics and their applications are described, followed by a survey of the current state of the literature regarding bioconjugation and ligation strategies.

Chapter 2 focuses on the design of tertiary amide-based conformational switches actuated through disulfide formation. The sequence Ac-GWPQG-NH_2 was studied through a double mutant cycle, making use of nuclear magnetic resonance (NMR) to evaluate the cis Pro content of the sequences and the cooperativity between the Trp and Gln residues. This information was then used to design two generations of disulfide-actuated switches, the conformations of which were studied through NMR.

Chapter 3 explores the design of novel β-turn mimics installed through chemical ligation. Hydrazide formation was employed in the design of three mimics, which are synthesised through a reductive amination step that allows the use of two unprotected peptide partners and mild conditions. The mimics were implemented in tripeptides and within the Tryptophan Zipper β-hairpin. NMR and circular dichroism (CD) analysis showed that the tripeptide systems adopted a turn-like conformation in solution, and that the full-sized conjugates successfully mimicked the β-turn and β-hairpin conformation in the control system.

Chapter 4 details the design, synthesis and evaluation of a cyclic peptide inhibitor of the MDM2/p53 protein-protein interaction, which contains one of the β-turn mimics designed in Chapter 3 and is synthesised by means of a ligation step that cyclises the peptide. The activity of the peptide was evaluated through surface plasmon resonance (SPR) assays, which showed an improvement in comparison to the control system, which contains a (D-Pro)-(L-Pro) β-turn.

Chapter 5 explores oxime ligation as a strategy to install β-turn mimics within β-hairpin backbones. One of the targets explored contained a β-turn mimic based on a *meta*-substituted aromatic ring, which afforded one stereoisomer that adopted a largely unfolded conformation, while the other showed successful mimicry of the turn and hairpin structure through CD analysis. An aliphatic oxime linkage was also explored, and although it provided less preorganisation it gave rise to two stereoisomers that successfully mimicked the conformation of the control system. The reduction of the oxime to generate an *N*-alkoxy amine junction was also studied through CD, which showed a higher degree of folding and stability in comparison to its oxime precursor.

Table of contents

Acknowledgements

I would firstly like to thank Dr Drew Thomson for giving me the opportunity to be the first PhD student in his group and to work on such interesting projects. I have enjoyed every second of my PhD and I have learned so much more than I expected, both about science and my passion for it, and about myself. Thank you for your help and support all along, both about science and my future endeavours.

Secondly, I would like to thank Dr Albane Neckebroeck for her help on the TrpZip project. Figuring out all those NMR conditions, impurities and assignments was a daunting task at first, but working with you made it feel doable. It was incredibly helpful to work side by side and of course good fun! I also want to thank Dr Brian Smith for all his help – learning peptide NMR was a big learning curve for me and he was always there to answer all my many questions, not only on this project but throughout my PhD. Thank you as well to Dr Sharon Kelly and Mrs June Southall for their help with CD, and to Stuart in IT for magically solving all my computer problems.

Thank you to everyone in the Chemical Biology lab, past and present. Working here has been a wonderful experience. Although this pandemic has made us miss quite a few, those Friday lab drinks really kept me going! Thank you to Alin, Tom and Roderick for making my first months in the lab so welcoming and for always keeping up the good vibes. Thank you to Susan for keeping the lab running, and to Alex for always being so cheerful and for his immense help. I learned so much from you. And of course thank you to the members of the Thomson group, it has been great working with you all!

Special thanks to Bebe Laura, Bebe Dee (potato gang!) and Kopano, I wouldn't be here without you guys. Your advice, support, coffee breaks and sometimes very necessary rants have been invaluable. In particular thank you Laura for your help and advice towards the end of my PhD, especially for reading my thesis. I of course need to thank Sarah and our child the salamander, although he didn't go as planned working with you always cracked me up. Also thank you to the late evening work team (Kopano and Sarah) – my motivation thanks you immensely. You turned work into a great time, and all your organic chemistry help was essential!

Thank you to Dem girls Emma and Lucy for being in this journey with me from day one, our lunch time meetings have become so much more! Your support and insight related to our PhDs but mostly in life are invaluable. I couldn't have done this without you two.

Ally, gràcies per ser al meu costat durant tota aquesta etapa, no me'n hauria sortit sense els nostres missatges de veu de 10 minuts i tot el teu suport. Encara que porti fora gairebé quatre anys, gràcies a tu sempre m'he sentit més a prop de casa. Thank you also to my Italian Nenas, Francesca and Giulia – although corona kept us apart lately, being in the city where we met has made me feel closer to you. Thank you for your emotional support even when you didn't know what I was working on!

Siempre queda de última, pero es la parte más importante: gracias a mi familia. Gracias mamá, papá y Sabela por vuestros consejos y apoyo durante esta etapa que ha traído alegrías y buenas noticias, aunque también altibajos y decisiones complicadas. Ha sido difícil estar lejos de vosotros durante este tiempo, pero a pesar de la distancia siempre me habéis hecho sentir acompañada, segura y capaz de todo. Gracias por escuchar mis innumerables dilemas y preocupaciones y por siempre ayudarme a encontrar mi camino.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, that the substance of this thesis is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

A portion of the work described herein has been published elsewhere as listed below:

Crecente-Garcia, S.; Neckebroeck, A.; Clark, J. S.; Smith, B. O.; Thomson, A. R. *Org. Lett.* **2020**, *22*, 4424–4428.

Selma Crecente García

Abbreviations

2-Abz	2-Aminobenzoic acid
Ac	Acetyl
Aib	α-Aminoisobutyric acid
Ala/A	Alanine
All	Allyl
Ar	Aromatic
Arg/R	Arginine
Asn/N	Asparagine
Asp/D	Aspartic acid
BME	β-Mercaptoethanol
Boc	*tert*-Butyloxycarbonyl
BOP	Benzotriazol-l-yl-oxy-tris-(dimethylamino)phosphonium hexa-fluorophosphate
BTM	β-Turn mimic
CCPN	Collaborative computing project for NMR
CD	Circular dichroism
CITU	Tetrachloro-*N*-hydroxyphthalimide tetramethyluronium hexafluorophosphate
COSY	Correlation spectroscopy
Cys/C	Cysteine
DBU	1,8-Diazabicyclo(5.4.0)undec-7-ene
DCC	*N*,*N*'-Dicyclohexylcarbodiimide
DCE	Dichloroethane
DCM	Dichloromethane
DIC	*N*,*N*′-Diisopropylcarbodiimide
DIPEA	*N*,*N*-Diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMC	Double mutant cycle
DMF	*N*,*N*-Dimethylformamide
DMS	Dimethyl sulfate
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid

DPDS	Dipyridyl disulfide
DSS	3-(Trimethylsilyl)-1-propanesulfonic acid-d_6 sodium salt
EDCI	*N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide
EDG	Electron donating group
EEDQ	*N*-Ethoxycarbonyl-2-ethoxy-1,2-dihydroquinoline
ESI	Electrospray ionisation
EWG	Electron withdrawing group
Fmoc	9-Fluorenylmethoxycarbonyl
FRET	Förster resonance energy transfer
Gln/Q	Glutamine
Glu/E	Glutamic acid
Gly/G	Glycine
GST	Glutathione S-transferase
Hao	2-[3-[Hydrazinecarbonyl]-4-methoxyanilino]-2-oxoacetic acid
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
HBA	Hydrogen bond acceptor
HBD	Hydrogen bond donor
HBS	Hydrogen-bonding surrogate
HCTU	O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HDM2	Human double minute 2
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
His/H	Histidine
HOAt	1-Hydroxy-7-azabenzotriazole
HOBt	Hydroxybenzotriazole
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear single quantum coherence
HT	High tension
IC_{50}	Half maximal inhibitory concentration
Ile/I	Isoleucine
IPA	Isopropyl alcohol
^{i}Pr	Isopropyl
k_a	Association rate constant
k_d	Dissociation rate constant

K_D	Equilibrium dissociation constant
LCMS	Liquid chromatography mass spectrometry
Leu/L	Leucine
Lys/K	Lysine
m/z	Mass to charge ratio
MD	Molecular dynamics
MDM2	Murine double minute 2
Met/M	Methionine
MRE	Mean residue ellipticity
NAMFIS	NMR analysis of molecular flexibility in solution
NCL	Native chemical ligation
NCS	*N*-Chlorosuccinimide
NMP	*N*-Methyl-2-pyrrolidone
NMR	Nuclear magenetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
NPT	Ensemble with constant pressure and temperature
NVT	Ensemble with constant volume and temperature
o/n	Overnight
o-NBS	*ortho*-Nitrobenzenesulfonyl chloride
PBS	Phosphate buffered saline
PDB	Protein data bank
PEG	Polyethylene glycol
Phe/F	Phenylalanine
Piv	Pivaloyl
p-PDA	*p*-Phenylenediamine
PPI	Protein-protein interaction
Pro/P	Proline
Proc	Propargyloxycarbonyl
p-TSA	*p*-Toluenesulfonic acid
PyBOP	(Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
ROE	Rotating frame Overhauser effect
ROESY	Rotating frame Overhauser effect spectroscopy
RP-HPLC	Reverse phase high pressure liquid chromatography

rt	Room temperature
SD	Standard deviation
Ser/S	Serine
S_N2	Bimolecular nucleophilic substitution
SPPS	Solid phase peptide synthesis
SPR	Surface plasmon resonance
^{t}Bu	*tert*-Butyl
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
Thr/T	Threonine
TIPS	Triisopropylsilane
T_m	Melting temperature
TOCSY	Total correlation spectroscopy
Trp/W	Tryptophan
TrpZip	Tryptophan Zipper
Trt	Trityl
TSP	Trimethylsilylpropanoic acid
Tyr/Y	Tyrosine
UV	Ultraviolet
Val/V	Valine

1 Introduction

1.1 Numbering and naming conventions

The residues in a peptide sequence are numbered from *N*- to *C*-terminus. Where a specific residue is considered the reference point, it is referred to as *i*. Subsequent residues are numbered *i+1*, *i+2*, etc., and preceding residues are numbered *i-1*, *i-2*, etc.

Undefined residues are referred to as Xaa, Yaa. Positions following and preceding a residue Xaa are referred to as Xaa+1 or Xaa-1, respectively.

The nomenclature recommended by the International Union of Pure and Applied Chemistry[1] has been adopted to identify the atoms in each residue for NMR analysis. Relevant examples are presented in Figure 1.

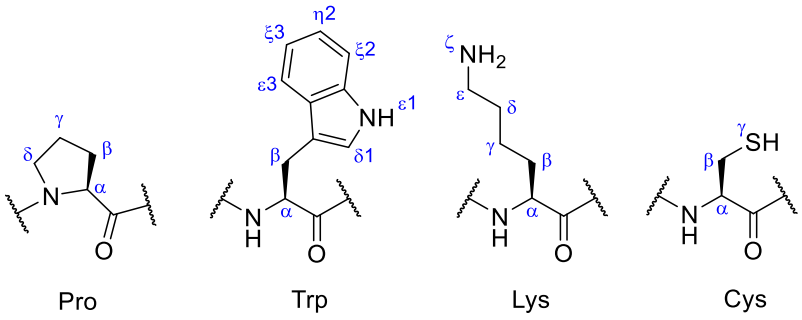


Figure 1. Nomenclature of relevant residues for NMR analysis.

1.2 Forces determining protein structure

The stability of a protein's three-dimensional structure is defined by the difference between the free energies of the folded and unfolded states (Equation 1). The native protein will be stable if the free energy necessary to unfold the protein (ΔG_U) is positive.[2]

$$\Delta G_U = G_u - G_f$$

Equation 1. Change in free energy necessary to unfold a native protein (ΔG_U), where G_u is the free energy of the unfolded state and G_f that of the folded state.

The individual free energies are composed of enthalpic and entropic terms (Equation 2), the latter being the major force favouring the unfolded state, as it provides more conformational freedom than the native folded protein.

$$\Delta G = \Delta H - T\Delta S$$

Equation 2. Change in Gibbs free energy, where H is enthalpy, T the temperature and S the entropy of the studied system.

A protein's folded state is only marginally more stable than the unfolded protein, with the ΔG of unfolding typically being 20 to 80 kJ/mol.[3] Folding is a cooperative process that takes place through the formation of many non-covalent interactions that are weak individually, but their summation creates the necessary difference in free energy so that the folded state will be favoured.

In this section, the forces that determine protein structure will be discussed, from the amide bonds that form the primary sequence and the post-translational disulfide bridges, to the non-covalent interactions that favour the folded state.

1.2.1 Thermodynamics: from folding to switching

Protein folding is a cooperative process, which means that intermediate conformations are not substantially populated and there are two distinct states at equilibrium (Figure 2).[4] The equilibrium constant for the folding process can be expressed as Equation 3, using the concentrations of the native and unfolded states that are present.

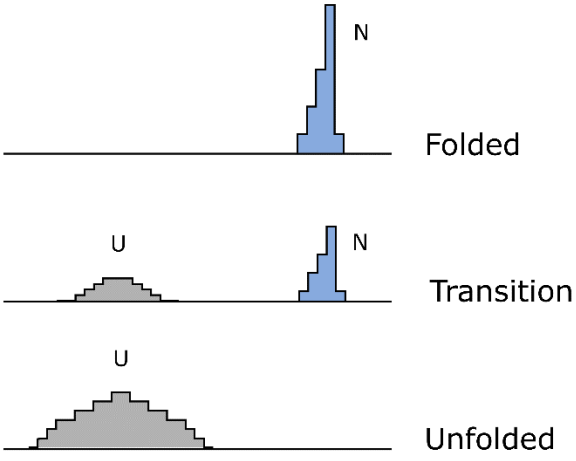


Figure 2. Co-operative, two-state folding transition between the native (N) and unfolded (U) states. Adapted with permission from ref.[4]

$$K_{eq} = \frac{[F]}{[U]}$$

Equation 3. Equilibrium constant for the folding process, where [F] is the concentration of the native folded state and [U] that of the unfolded state.

The stability of the folded native protein (ΔG°_{F}, Equation 4) can be measured from the free energies of the folded and unfolded states:

$$\Delta G^{\circ}_{F} = G^{\circ}_{F} - G^{\circ}_{U} = RTlnK_{eq}$$

Equation 4. Net stability of the folded state (ΔG°_{F}), where G°_{F} is the free energy of the native state, G°_{U} that of the unfolded state, R is the gas constant (8.31 $JK^{-1}mol^{-1}$) and T the temperature (K).

The same analysis can be applied to two-state systems, for example the binding or unbinding of a substrate, two distinct folded conformations as well as a two-state switch. In the latter case, for example a peptide with *cis* and *trans* configurations, the population of said states at equilibrium is defined by the Boltzmann distribution (Equation 5), which dictates that the most populated state will be the one with the lowest energy and therefore most stable.

$$\frac{N_A}{N_B} = e^{-\left(\frac{\Delta\Delta G}{RT}\right)}$$

Equation 5. Boltzmann distribution, where N_A and N_B are the number of particles in each state, R is the gas constant, 8.31 $JK^{-1}mol^{-1}$, T is the temperature in K and $\Delta\Delta G$ the Gibbs free energy in J.

If $\Delta G \sim k_B T$, the energetic difference between both states is small and the population will be evenly distributed between them. Therefore, the distribution of particles is defined by the Gibbs free energy, which can be related to the equilibrium constant between states through Equation 6:

$$\Delta G^{\circ} = -RTlnK_{eq}$$

Equation 6. Relationship between $\Delta G°$ and the equilibrium constant (K_{eq}), where R is the gas constant, 8.314 $JK^{-1}mol^{-1}$, and T is the temperature in K.

In order for switching to take place, enough energy needs to be provided to overcome the energy difference between states and invert the distribution of the population. For example, to invert a population from 9:1 to 1:9, the difference in energy is 5.4 kJ/mol. Therefore an energy of 10.8 kJ/mol needs to be provided for the population inversion to take place. In the case of switching from a *trans* peptide bond to the *cis* state, the $\Delta\Delta G$ between states is approximately 10 kJ/mol (Figure 3).[5] The preference for one state over another can be biased through the introduction of a new feature into the system[6] – for example, the formation of a disulfide bridge or the phosphorylation of a residue could provide enough energy for a peptide to switch to the *cis* configuration.

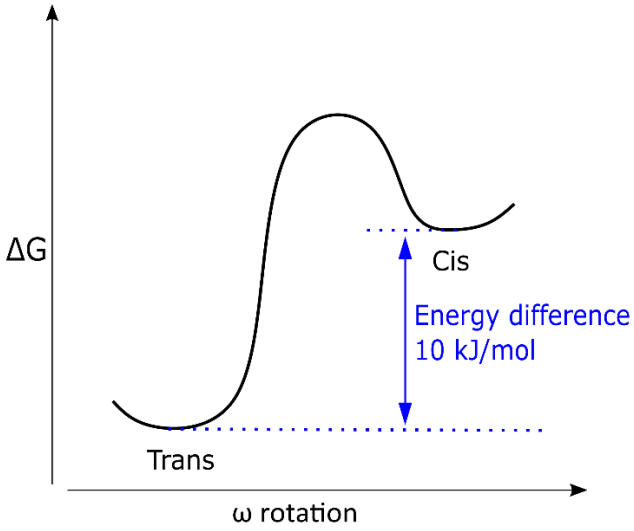


Figure 3. General energy diagram for the isomerisation of a peptide bond. The typical difference in energy between the *cis* and *trans* states is represented in blue.

The strength of non-covalent pairwise interactions in peptides can be measured through double mutant cycles (DMC), utilising the relationship between the population of states and the Gibbs free energy.[7] DMCs are based on introducing two mutations into the system under study, which creates a thermodynamic cycle that allows for the calculation of the free energy of interaction between residues. The ΔG associated to a structural property of the protein can be calculated at every step of the cycle, therefore if the ΔG upon a double mutation is different than the sum of the ΔGs after single mutations, the residues are coupled.[8] These mutations can be a change in residues as well as the protonation of the system or a change in conditions such as pH.

The introduction of more than one mutation is necessary in order to cancel the secondary effects caused by the interaction of a residue with other parts of the protein. These cycles also require a two-state folding behaviour, that is a distinct folded and unfolded state, or two distinct conformations in order to use them in folding systems like peptides.[9]

DMCs have been widely used to measure the strength of non-covalent interactions between residues in peptides: from cross-strand interactions in beta hairpins[10,11] to quantifying electrostatic[12] and hydrophobic interactions.[13] Focusing on the mutation of two residues in a peptide system, the DMC can be expressed as:

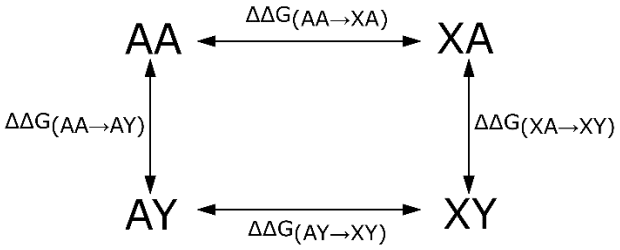


Figure 4. General double mutant cycle showing the ΔΔG for each mutation.

Usually the reference state consists of the residues studied being replaced by Ala in order to minimise the formation of new side chain interactions. Therefore, the ΔΔG of the interaction between residues X and Y can be expressed as:[8]

$$\Delta\Delta G_{int} = \Delta\Delta G_{AA\rightarrow XA} - \Delta\Delta G_{AY\rightarrow XY} = \Delta\Delta G_{AA\rightarrow AY} - \Delta\Delta G_{XA\rightarrow XY}$$

Equation 7. Calculation of ΔΔG of the interaction between both mutated residues.

Where the free energy for each mutation corresponds to the subtraction of the ΔG of both states, therefore:

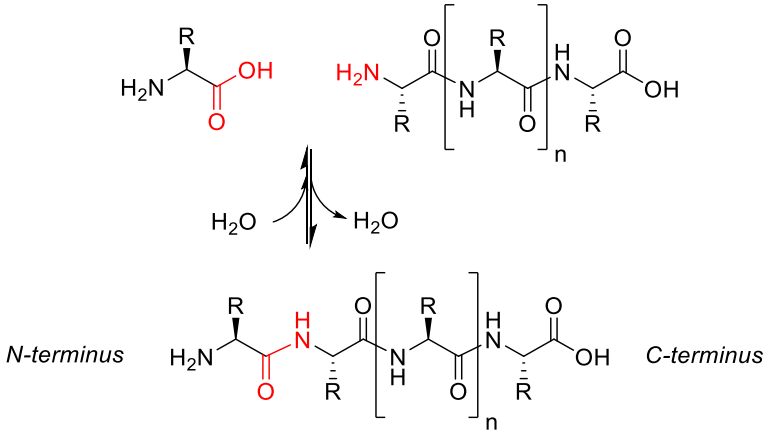
$$\Delta\Delta G_{int} = \Delta G_{AA} - \Delta G_{XA} - \Delta G_{AY} + \Delta G_{XY}$$

Equation 8. Calculation of the ΔΔG of interaction using the individual ΔG terms for each mutant.

$\Delta\Delta G_{int}$ measures the coupling between both residues studied: if its value is not zero, then the residues are coupled. If $\Delta\Delta G_{int}$ is positive, negative cooperativity is taking place and if it is negative, the residues have positive cooperativity in stabilising the same conformation.[14]

1.2.2 The amide bond

Proteins and peptides are composed of amino acids linked through peptide bonds. These bonds can be formed through a condensation reaction between the carbonyl carbon of the residue being added to the chain and the NH of the following residue (Scheme 1), but under standard biochemical conditions and at physiological pH this reaction does not occur, as the hydroxyl group is a poor leaving group.[15]



Scheme 1. Condensation reaction that generates an amide bond.

In cells, proteins are synthesised through transcription and translation from DNA, but smaller-sized peptides can be chemically synthesised. In a synthetic setting, peptides can be made through the activation of the acid by introducing an electron withdrawing group on the acyl carbon prior to the attack by the amino group (Figure 5). Many methods have been developed for this purpose,[16,17] which involve the use of diverse coupling reagents sometimes in combination with additives in order to prevent epimerisation.

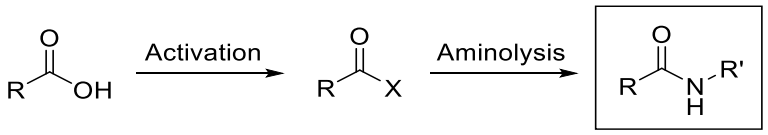


Figure 5. General pathway for peptide coupling.

The hydrolysis of a peptide bond is worth from -2 to -17 kJ/mol in free energy;[18] this value depends on which specific amino acids are involved as well as the neighbouring amino acids and the pH of the solution.

The delocalisation of the lone pair of the nitrogen atom that takes place along the peptide bond confers a partial double bond character to it (Figure 6). This causes the rotation about the amide bond to be restricted and therefore the C^{α} and C=O of residue Xaa as well as the NH and C^{α} of residue Xaa+1 are coplanar. Another consequence of resonance is the length of the peptide bond (on average 1.32 Å), which is in between that of a C-N bond (1.45 Å) and a C=N bond (1.25 Å).[3]

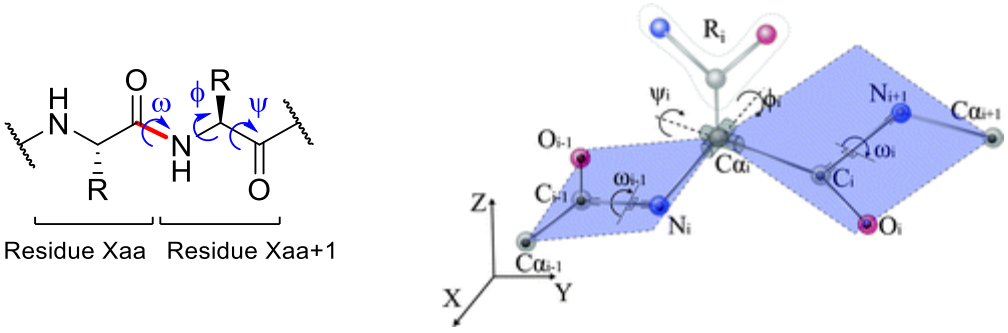


Figure 6. Left: peptide chain showing the dihedral angles in blue. The amide bond is represented in red. Right: planes that describe the dihedral angles. Reproduced with permission from ref.[19]

1.2.3 Disulfide bridges

Disulfide bridges are the second most common covalent link between two residues in proteins, which form through an oxidation reaction between the sulfur atoms in two cysteine residues (Scheme 2). The redox potential for this reaction in proteins is of -0.25 V.[20] Disulfide bridges often take place between residues widely separated in the primary sequence, although local ones are also observed linking loops with positive β-turn potential.[21]

$$R_1\text{-SH} + R_2\text{-SH} \rightleftharpoons R_1\text{-S}-\text{S-}R_2 + 2H^+ + 2e^-$$

Scheme 2. Formation of a disulfide bridge.

Disulfides are comprised of the C^{α}-C^{β}-S^{γ}-S'^{γ}-C'^{β}-C'^{α} atoms, where the S-S distance is 2 Å,[22] and they are defined by the dihedral angles χ1, χ2 and χ3 (Figure 7). Rotation can take place about the S-S bond (the χ3 dihedral angle), which defines two types of disulfides: left-handed, with an average of χ3 = -87°, and right-handed, with χ3 = +97°.[23]

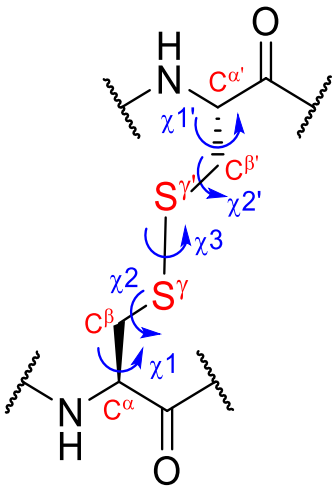


Figure 7. Disulfide bridge showing the dihedral angles in blue and the relevant atoms in red.

The folding kinetics of proteins are affected by the location of the disulfide bridge relative to the folding nucleus (the folding nucleus is a region of the transition state where specific residues have come into contact so that the protein has overcome its free energy barrier, and will rapidly fold into its native conformation[24]). If the bridge is near or in the folding nucleus, it will stabilise the intermediate folding states and therefore accelerate folding. But if it connects elements of secondary structure that fold late in the process, the disulfide can slow folding considerably and it will not provide additional stability until the protein is completely folded.[25,26] This was applied in a β-hairpin as a strategy to populate a folding transition state in order to study the folding pathway and kinetics of the system.[27] The authors found that including a disulfide crosslink near the β-turn region increased the folding rate by an order of magnitude.

Disulfides are also thermodynamically coupled to the folding process – studies by Tidor *et al.*[28] as well as Pace *et al.*[29] showed that each bridge contributes 10.5 kJ/mol to 21.3 kJ/mol to the thermodynamic stability of the protein. They stabilise the folded conformation by reducing the conformational fluctuations of the unfolded state, therefore decreasing its entropy and making it less favourable compared to the native folded protein.[21] The degree of stabilisation provided by a disulfide bridge depends on two factors: the size of the loop created and the location of the bridge. The increase in stability provided to the native structure is directly proportional to the number of residues between the bridged cysteines,[26,30] and, as found by Matsumura *et al.*,[31] a disulfide will be most stabilising if it is located in a flexible region.

Disulfide bridges are required for folding and stability, which directly impacts the activity of many proteins. Narhi *et al.*[32] studied a growth factor protein which contains three disulfides and synthesised mutated analogues that lacked one, two or all bridges. They found that if all disulfides were removed, the protein was unfolded and became inactive. If one or two of the disulfide bridges were present, a compact but partially folded state was formed and some biological activity was observed, although lower than in the native protein.

Disulfides can also be directly implicated in biological functions[30] and are classified as catalytic and allosteric. Catalytic disulfides are present in enzyme active sites like oxidoreductases and they participate in many cysteine-based redox-systems, including thiol-disulfide exchange reactions, which maintain the redox balance in cells through the glutathione/glutathione disulfide redox-couple.[33,34] Allosteric disulfides[35] regulate protein function by triggering conformational changes. They are labile disulfides that have high torsional strain as well as a more stretched sulfur-sulfur bond compared to non-labile ones: database studies[22] have shown that the mean S-S bond length for all disulfides in the PDB is 2.046 Å, whereas it is 2.055 Å for labile disulfide bridges.

Disulfide bonds are a widely used tool in protein engineering[36] as well as in the development of new peptide-based therapeutics. Cyclisation using this strategy can lock the peptide into its bioactive conformation and thus improve its potency and selectivity.[30]

1.2.3.1 Vicinal disulfides

In early studies vicinal disulfides were thought to belong to a list of 'forbidden' regions which could not form a disulfide bond. These regions were described by Thornton[21] in 1981 and they included regions both in the primary and secondary structure which generated too much strain in the protein conformation. Nowadays it has become clear that these regions are indeed populated by disulfide-bridged cysteines[37] and although rare, vicinal disulfides do occur in proteins.

Vicinal disulfides are formed between two adjacent cysteine residues in the primary sequence. Although there is a strong preference for a *trans* peptide bond between the cysteines involved in a vicinal disulfide, *cis* peptide bonds are also observed (Figure 8). The formation of the disulfide bridge causes significant distortion to the backbone conformation, thus a *trans* Cys peptide bond adopts an average ω angle of 171°. *Cis* bonds suffer less distortion and usually have ω close to 0°.[38]

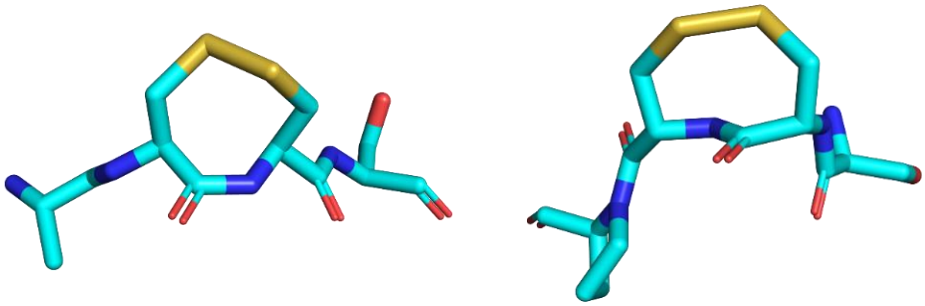


Figure 8. Vicinal disulfide bridges with *cis* (left, PDB 3HOL) and *trans* (right, PDB 2QC1) Cys.

Although they are scarce, vicinal disulfides have been found to be conserved motifs in proteins and they are usually located in tightly packed domain interiors or in relatively mobile loops.[39] Many of them undertake structural roles and some examples have been shown to be necessary for activity,[40] including their participation in redox reactions.[37] An interesting example is the Microcin C7 self-immunity protein (MccF), a serine peptidase that cleaves linkages between amino acids and nucleic acids and is responsible for breaking down the antibiotic Microcin C. MccF has a catalytic triad composed of Ser-His-Glu which is preceded by two consecutive Cys residues. In its active state, the Cys in MccF are in the reduced thiols form, but upon formation of a vicinal disulfide between them, with a *cis C*-terminal Cys, the backbone conformation is altered significantly. This causes MccF to become inactive.[39]

The frequent reduction of vicinal disulfides and oxidation of the free thiols as well as the pronounced difference in conformation between both states has led to an interest in whether these systems could be used as a switch. This has been observed in naturally occurring peptides,[41] but no designed systems have been yet reported.

1.2.4 Van der Waals interactions

Van der Waals interactions arise from the fluctuation of the charge density in an atom's electron cloud, which creates a temporary dipole in the atom. If another atom is close enough, that dipole will induce an opposite transient dipole in the nearby atom, and a weak attraction will take place. As both partners get closer together, the interaction will become stronger until they reach the Van der Waals contact distance (Figure 9). At this point, the attraction between dipoles is balanced by the repulsion between the atoms' electron clouds. The Van der Waals radius of an atom defines how close another atom can approach it before experiencing repulsion.[15]

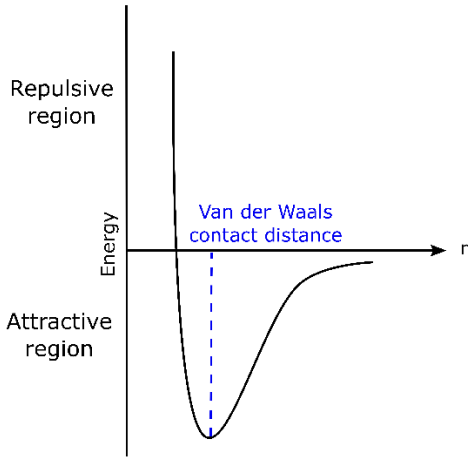


Figure 9. Diagram of the energy of a Van der Waals interaction as a function of the interatomic distance.

The energy of these interactions is described by the Lennard-Jones potential, and it depends on the difference between the attractive (r^{-6}) and the repulsive (r^{-12}) terms (Equation 9). The energy for this interaction per atom pair usually ranges from 2 to 4 kJ/mol.[42]

$$E_{VdW} = \sum E \left[\left(\frac{r_m}{r} \right)^{12} - 2 \left(\frac{r_m}{r} \right)^{6} \right]$$

Equation 9. Energy of a Van der Waals interaction (E_{VdW}), where r_m is is the minimum energy interaction distance, E is the depth of the potential well and r is the interatomic distance.

The repulsive term refers to the electronic repulsion between the interacting atoms and the attractive term encompasses three different contributions: the strongest is the orientation effect, or the interaction between permanent dipoles, which depends on the relative orientation between the participating atoms. The induction effect is the interaction between permanent and temporary dipoles and it varies depending on the polarisability of the molecules involved. Lastly, the London force or dispersion effect, which refers to the constant shifting of dipoles and arises from the induction of a complementary dipole in a nearby molecule by a temporary dipole.[3]

1.2.5 Electrostatic interactions

Electrostatic interactions arise from the attraction or repulsion between charged species. There are many types depending on the groups involved, as they can take place between permanently charged atoms, dipoles and quadrupoles. They will all be introduced in the following sections.

Electrostatic interactions are defined by Coulomb's law (Equation 10) and they are the non-covalent forces that act over the longest distance, as can be observed by their distance dependence: for the classic ion-ion interactions the distance dependence is of r^{-1}, whilst for ion-dipole interactions it is of r^{-2}.

$$E = \frac{Q_1 * Q_2}{4\pi\varepsilon * r}$$

Equation 10. Coulomb's law for an ion-ion interaction, where Q_1 and Q_2 are the magnitudes of the charges, ε is the electric constant and r the distance between the charges.

These interactions can play a role in protein stability by the establishment of specific charge interactions (or salt bridges, discussed in the next section) and through non-specific long-range interactions.[43] In the latter case, repulsions between partially charged atoms can take place. In proteins the groups that participate in electrostatic interactions can be residue side chains as well as the terminal groups. These are usually found on the protein surface, where their charges can interact with the solvent due to solvation.[44] The net charge of a protein also plays a role in stability: if a protein is in an extreme pH environment, the native protein can be destabilised due to the greater charge density present in the folded protein over the unfolded state.[43]

Electrostatic interactions play a key role in promoting the right-handed twist in β-hairpins. Although there is an intrinsic bias for a right-handed twist that arises from the residue side chains steric hindrance and intra-strand interactions,[45] that is an energetically small contribution. The dominant driving forces for the twist are inter-chain interactions that are mainly electrostatic and Van der Waals forces.[46]

1.2.5.1 Ion-ion interactions: salt bridges

Ion pairing interactions take place between permanently charged molecules or ions, for example Na^+ and Cl^-. This Coulombic attraction can also take place between charged residues in proteins, in which case they are referred to as salt bridges.

Salt bridges are a pH-dependent, attractive interaction that takes place between oppositely charged residues that are within ~5 Å of each other (Figure 10). The negatively charged side chain can be Asp, Glu, Tyr, Cys and the *C*-terminal carboxylate group and the positively charged partner can be Lys, Arg, His and the *N*-terminal amino group. These interactions

can be found buried within the protein core or on the surface, they can be accompanied by a hydrogen bond between the partners, and they can also be networked, in which case at least one of the participating residues establishes another salt bridge with a different residue in the protein. In a database study, Kumar *et al.*[47] found that 34% of all salt bridges were buried, 8% of them were networked, and that most of them were formed between residues close in the primary sequence.

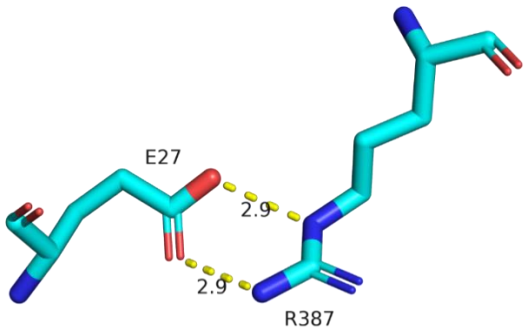


Figure 10. The most stabilizing salt bridge in the database by Kumar *et al.*[47] between E27 and R387 in the PDB accession code 1SMD. The distances between atoms are shown in Å.

Many approaches to calculate the contribution of salt bridges to a protein's stability have been designed,[2] which have shown that salt bridges contribute from 4.2 kJ/mol to 16.7 kJ/mol in free energy[48] depending on whether they are found on the protein's surface or if they are buried, respectively.

Salt bridges are composed of two energetic contributions: the Coulombic interaction between opposite charges constitutes the direct effect, which is a favourable contribution. This attraction is counterbalanced by the indirect effects, which tend to be unfavourable and include the desolvation cost and the background interactions between the charges and dipoles in their immediate environment. This balance between contributions defines whether a salt bridge will contribute to folding or destabilise a conformation.[2] In general, there is a large desolvation cost but this is compensated for by the stronger electrostatic interactions that take place within the protein core: in a database study, Kumar *et al.*[47] showed that 85.6% of all salt bridges were stabilising. Kumar *et al.*[47] also found that 86% of all salt bridges contained a hydrogen bond between their side chains, highlighting how hydrogen-bonded salt bridges establish stronger interactions and contribute towards folding.

Salt bridges have been shown to contribute to the stability of α-helices[49] and β-hairpins.[13] For example, Ciani *et al.*[50] substituted two cross-strand residue pairs in a β-hairpin by Glu-Lys salt bridges and found that the individual mutations contributed -1.2 to -1.3 kJ/mol to

hairpin stability. The introduction of both mutations at the same time contributed -3.6 kJ/mol to the free energy, which showed a cooperative behaviour between them.

1.2.5.2 Ion-dipole interactions

Ion-dipole interactions take place between a charged atom and another which bears a permanent dipole. A classic example is solvation, or the interaction between ions like sodium chloride and water (Figure 11). The strength of these interactions varies depending on the magnitude of the dipole and charge involved as well as their relative orientation.

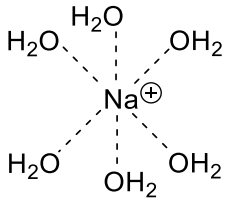


Figure 11. Example of an ion-dipole interaction: solvation of a sodium ion in water.

For these interactions to take place in proteins, the ion has to be completely desolvated and buried in the protein core and it interacts only with neutral dipoles: the Ser, Thr, Asn, Gln, Tyr and Trp side chains as well as the backbone NH and carbonyl groups. The partner dipole often establishes other secondary interactions that favour a more rigid binding geometry.[51] The ion-dipole interaction encompasses more specific interactions like charge-neutral hydrogen bonds and coordination. In proteins these interactions play a wide variety of roles, from binding substrates to participating in catalysis or coiled-coil stabilisation.[51]

1.2.5.3 Dipole-dipole interactions

Hydrogen bonding

Hydrogen bonds arise from the electrostatic attraction between two dipoles, one which has a hydrogen atom covalently bound and acts as a hydrogen bond donor, and another with a partially negative charge that acts as the acceptor (Figure 12). The distance between heavy atoms in C=O···NH hydrogen bonds in proteins is 3 Å on average.[52] They are very directional: the strongest interaction takes place when the dipoles are aligned with each other, although misalignment is common in proteins due to strong geometrical constraints. The tolerated deviation is usually of +/- 40°.[3]

Figure 12. Hydrogen bond interaction (in red), showing the donor (in blue) and acceptor (in pink).

Hydrogen bonds are relatively weak, with energies for a main chain C=O···NH hydrogen bond of 15 to 20 kJ/mol.[44] In proteins, about 90% of all internal polar groups participate in hydrogen bonds,[53] which shows how they are key in determining protein structure. These interactions mainly take place between the main chain amide NH and carbonyl groups (Figure 13), but they can also occur between side chains, between a side chain and the main chain, as well as with water molecules within the protein core.

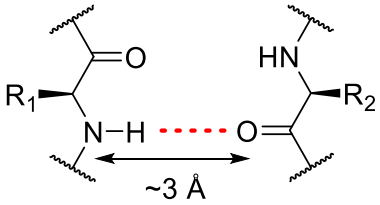


Figure 13. Hydrogen bond (in red) between two main chain amides, showing the distance between heavy atoms.

The energy of hydrogen bonds is not completely described by electrostatic interactions, as there is also the contribution of orbital interactions.[54] The formation of a hydrogen bond (X-H···Y) often causes an elongation of the X-H bond, and these interactions also show angular constraints that are greater than the expected for purely electrostatic interactions.[55] This suggests that an orbital overlap is taking place and that these bonds have a partial covalent character. This effect is caused by the donation of charge from the n orbital in the acceptor atom (Y) into the X-H σ* orbital, which causes an increase of the latter's antibonding character and results in a weakening and lengthening of the X-H bond.[56] Although this is seen as evidence for the covalent character of hydrogen bonds, a consensus on the amount of covalency in these interactions is still missing.[54]

α-Helices and β-sheets are greatly stabilised by hydrogen bonds. In α-helices, main chain hydrogen bonds that stabilise the helical conformation take place between the carbonyl oxygen of residue *i* and the amide NH group of residue *i+4*. They are generally ~3 Å between heavy atoms[3] and lie parallel to the helical axis (Figure 14).

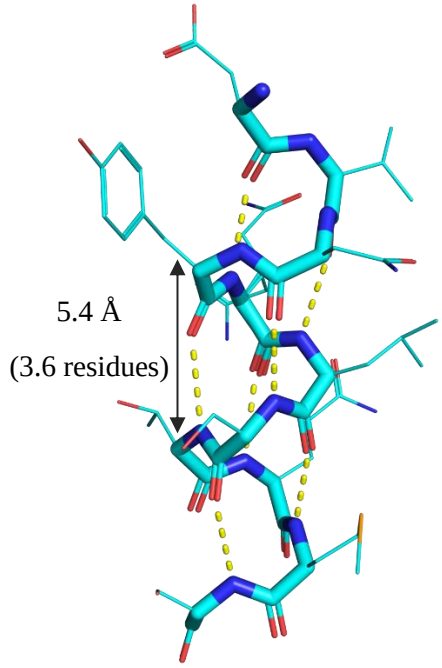


Figure 14. α-Helix structure, showing the backbone hydrogen bond pattern in yellow. PDB accession code 3IEE.

Adjacent β-strands in β-sheets are associated through inter-strand hydrogen bonds, which take place between the backbone carbonyl and NH groups. Strands can align in a parallel and antiparallel manner (Figure 15), each of which has a specific hydrogen-bonded pattern.

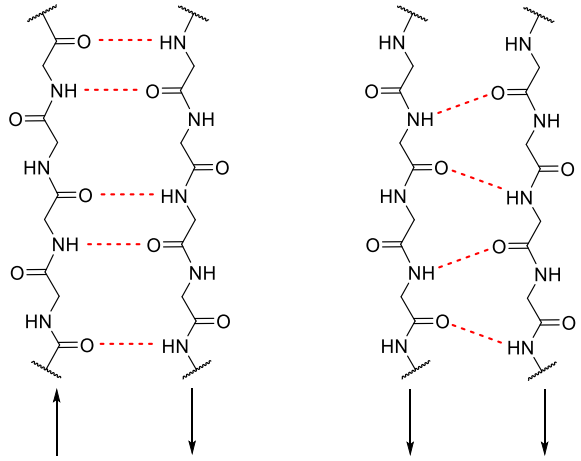


Figure 15. Antiparallel (left) and parallel (right) β-sheets, showing the backbone hydrogen bonding pattern in red lines.

n to π* interaction

The n to π* interaction takes place between a carbonyl group and a nucleophile, which is usually another carbonyl group. In this interaction, the carbonyl group acting as the nucleophile donates lone pair electron density into the other group's empty π* orbital (Figure 16).[57]

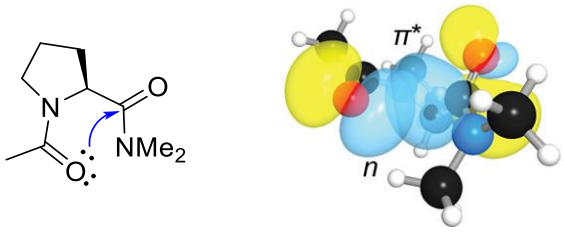


Figure 16. n to π* interaction between two carbonyl groups and overlap of the n and π* orbitals (in blue). Reproduced with permission from ref.[58]

Although this interaction was initially described as a purely dipole-dipole interaction, recent research has shown that is not its only component. It is composed of a Coulombic contribution between the partial charges in the carbonyl groups, the dipolar interaction between permanent dipoles, and a donor-acceptor interaction which involves the donation of electron density.[57]

The strength of the individual interactions is relatively weak, ranging from 1.3 to 2.9 kJ/mol,[57] but their overall impact in protein structure is significant due to carbonyl groups being ubiquitous. A search of the PDB performed by Bartlett *et al.*[59] showed that 34% of residues had the right orientation to participate in an n to π* interaction. The authors also performed DFT calculations to find out what dihedral angles give rise to these interactions and compared them to the angles from the residues found in the PDB. This confirmed that those 34% of residues engage in said interactions and highlights the relevance of this interaction in folded proteins.

The n to π* interaction between two carbonyl groups in a Xaa-Pro peptide bond is related to the conformation of the Pro residue, as the orbital overlap between the lone pair of the Xaa residue and the π* orbital can only take place when the peptide bond is in the *trans* configuration.[60]

1.2.5.4 Ion-quadrupole interactions

Two types of ion-quadrupole interactions can take place: the most widespread and most studied is the cation-π interaction, although in recent years the anion-π interaction has received more attention[61] and it has also been recognised as an important non-covalent interaction in protein structure.

Anion-π interactions take place between a negatively charged ion or residue and an electron-deficient π system. An analysis of the PDB by Lucas *et al.*[62] showed that 61.3% of all

structures present these interactions, and that the most common partners involved are Glu and His (16.7% of the interactions). Glu and Asp were the most frequent anionic partners, interacting through a parallel orientation between the carboxylate group and the π-system (Figure 17).

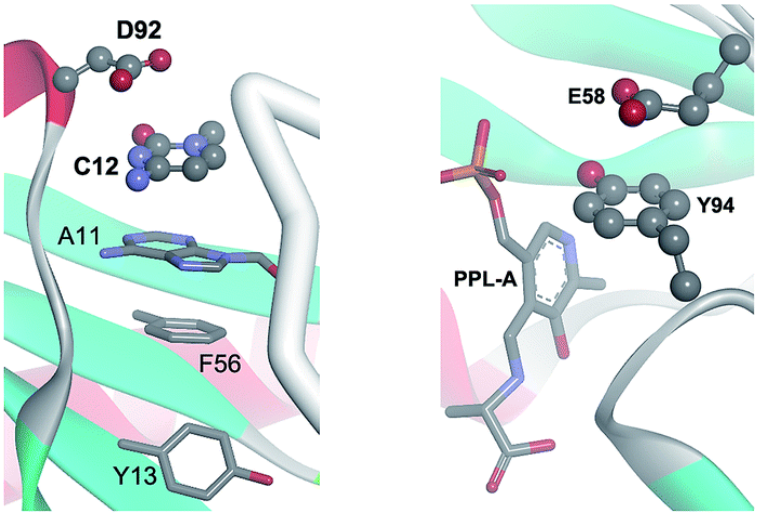


Figure 17. Left: Asp-π interaction (PDB accession code 1URN). Right: Glu-π interaction (PDB accession code 1WC7). Reproduced with permission from ref.[62]

Smith *et al.*[63] performed the mutation of two positions in a WW domain (small protein domain with two conserved Trp residues that folds into a triple-stranded β-sheet) that are brought close together by the reverse turn into Asp and Phe, and found that the anion-π interaction generated contributed -5.4 kJ/mol to the protein's stability.

Cation-π interactions take place between an aromatic side chain (Phe, Tyr, Trp or His) and a positively charged partner (Lys, Arg or protonated His), and arises from the electrostatic attraction between the cation and the electron-rich face of the quadrupole in the π-system (Figure 18).

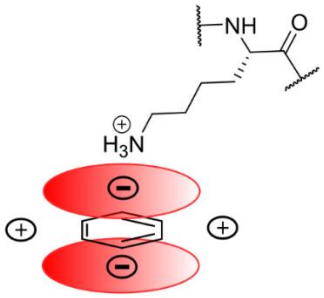


Figure 18. Cation-π interaction between lysine and a π-system, showing its quadrupole character.

Cation-π interactions were originally considered a weak interaction without a major role in folding, but in the late 1990s more studies were performed that recognised them as one of the most important non-covalent interactions in protein stability. Gallivan *et al.*[64] performed a database search to analyse the frequency with which each amino acid was involved in a cation-π interaction (His was excluded from this study). 70% of all Arg side chains were located near an aromatic residue, indicating a strong preference over Lys as the cationic partner. Trp was the preferred aromatic partner, with 26% of all Trp in the PDB being involved in an energetically relevant cation-π interaction, and Arg-Trp was the most commonly observed amino acid pair (40%).

Ma *et al.*[65] expanded the definition of the cation-π interaction as an 'amino-aromatic interaction' in biological systems and divided it into two classes: one where a positively charged side chain is involved (Arg, Lys and protonated His) and another with a neutral nitrogen-containing side chain (Asn and Gln). The first class is stronger (with interaction energies between -1.7 and 10.0 kJ/mol[66]) and can also take place through the CH_2 adjacent to the positively charged group, whereas the latter class establishes weaker interactions that can only take place through the NH_2 group.

Two kinds of geometries can take place: one with a perpendicular arrangement where the N-H points towards the face of the aromatic ring and another with a parallel arrangement (Figure 19).

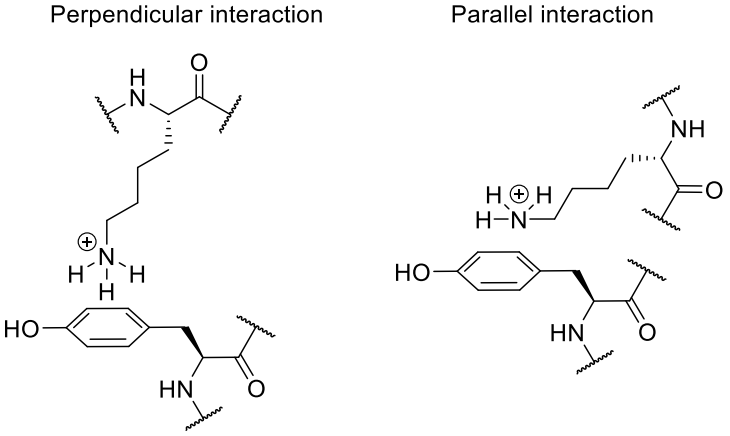


Figure 19. Possible geometries of a cation-π interaction between Lys and Tyr.

1.2.5.5 Dipole-quadrupole interactions: CH-π interactions

CH-π interactions are a weak hydrogen bond that takes place between a CH donor and an aromatic π-acceptor. Many groups can act as those partners, but the most prominent

interactions take place between an aliphatic or aromatic CH donor and an aromatic π-acceptor. In those cases, the energy of the interaction ranges from 6.3 to 10.5 kJ/mol.[67] They have a dual nature in that they are composed of two energetic terms:[68] the electrostatic component, which defines the orientation of the C-H bond towards the face of the aromatic ring, and a dispersion component, which is the major source of the attraction between both partners.[69]

Brandl *et al*.[70] examined the PDB to analyse the frequency with which these interactions occur, what types exist and residue preferences. They found that CH-π interactions are ubiquitous in proteins and classified them into 12 different classes depending on the donor/acceptor groups involved. On the donor side, the CH group can be aliphatic, aromatic or a C^{α}-H; the authors observed a preference for aromatic (Tyr $C^{\delta 2}$-H 4%) and aliphatic donors, some of the preferred residues being Lys (C^{γ}-H 3%), Met (C^{γ}-H 2.7%), Arg (C^{γ}-H 2.7%) and Pro (C^{δ}-H 3%, C^{β}-H 2.8%, C^{α}-H 2%). On the acceptor side, it can be an amide, carboxylate, guanidinium or aromatic side chain. A preference was observed for the latter class, with Trp being the most frequently observed (71.7% frequency, interacting through either ring). About 40% of the interactions were found to take place between residues separated by nine residues or fewer in the primary sequence.

The CH-π interaction has a key role in the stabilisation of protein secondary structures: in antiparallel β-sheets, Pro-aromatic and Gly-aromatic pairs are overrepresented[67] due to their involvement in cross-strand CH-π interactions (Figure 20). In the case of Pro, the H^{α} and H^{δ} have a partially positive charge,[71] which causes the polarisation of the CH bond that favours the interaction with the partially negative π-face of the aromatic ring.

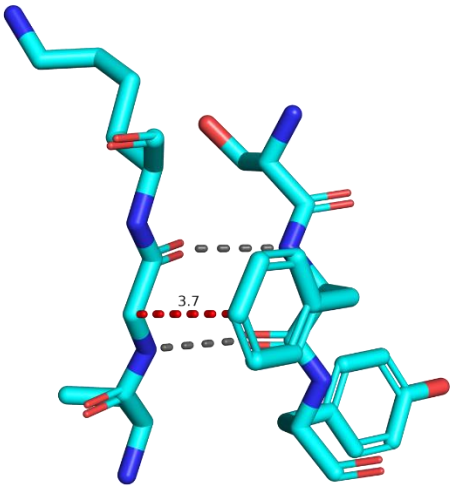


Figure 20. CH-π interaction between Gly and Phe (in red) along an antiparallel β-sheet. The hydrogen bonds are represented in grey.[67]

Aromatic residues are favoured in the *i+1* position in type VI β-turns, as they favour the occurrence of a *cis* Pro in the *i+2* position. This stabilisation of the reverse turn is mediated by a CH-π interaction between the H^{α} or H^{δ} atoms in the Pro ring and the preceding aromatic residue (Figure 21).

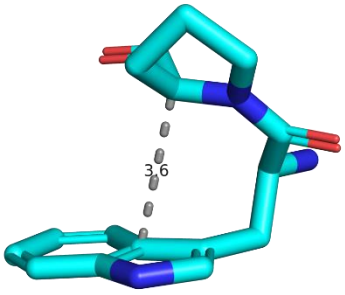


Figure 21. CH-π interaction between Trp and Pro (PDB accession code 1OAC).

1.2.5.6 Quadrupole-quadrupole interactions

The effect of quadrupole-quadrupole (or π-π) interactions in protein folding and stability can be considered through two components: the hydrophobic effect and local π-π interactions. These interactions are weak individually, but in conjunction they represent the most important contribution to the stability of the protein's folded state. In this section, π-π interactions are described, while the hydrophobic effect is discussed in Section 1.2.6.1.

Local π-π interactions can take place between a wide range of functional groups, using multiple points of contact and in multiple geometries. They are widespread in nature, with a classic example being DNA base pair stacking.[72] Their role in protein stabilisation was first analysed by Burley and Petsko,[73] who were the first to propose that aromatic-aromatic

interactions were the last non-bonding interaction key in protein folding. They found that the free energy of formation for each pair was -2.5 to -5.4 kJ/mol and proposed that these interactions could act as nucleation sites during folding, as they only found them in ordered regions of the protein. In antiparallel β-sheets, cross-strand hydrophobic side chain-side chain interactions are vital for sheet formation,[74] Trp-Trp pairs conferring the greatest stability to β-hairpins.[75]

The interaction between two aromatic systems is ubiquitous in proteins: on average, 60% of the aromatic residues in a protein participate in a π-π interaction.[66] The strength of the interaction was measured by Hunter *et al.*,[76] who have done a great amount of work in this area. They used the double mutant cycle approach to measure the magnitude of an edge-to-face interaction in a molecular zipper complex and found that if different substituents were introduced in the ring, the free energy was affected. The values for the interaction energy ranged from +1.0 kJ/mol to -4.9 kJ/mol, the unsubstituted interaction being worth –1.4 kJ/mol. Even though the magnitude of one π-π interaction is weaker than a hydrogen bond, the presence of many can have a considerable effect in the structure of a protein.

π-π interactions are made up of two components, namely electrostatic and Van der Waals interactions. The electrostatic component defines the orientation and geometry of the interaction, whilst the Van der Waals component is larger in magnitude and therefore the main source of attraction.[69,77]

There are very strong geometrical requirements for these interactions to be favourable, which arise from the fluctuation of the electrostatic attraction depending on the angle with which both rings stack and the offset between them. Hunter and Sanders[78] explored this and found that the face-to-face arrangement was unfavourable, whilst the edge-to-face geometry and the parallel stack with an offset were favoured conformations. Their results can be interpreted though the graph in Figure 22.

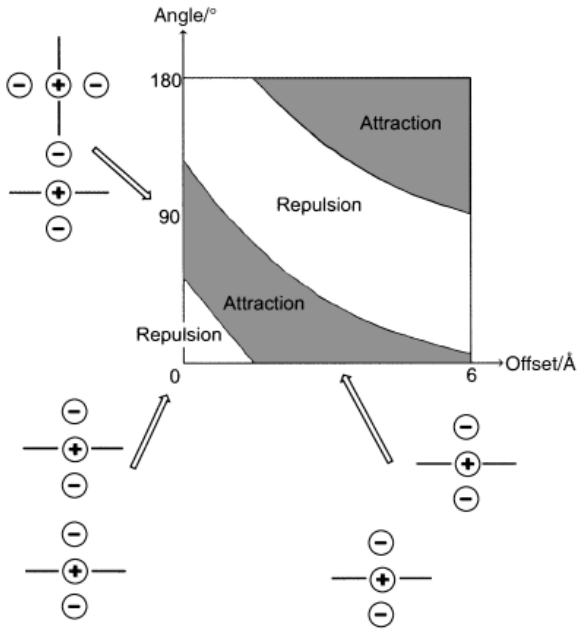


Figure 22. Orientation of the quadrupoles in π-systems as a function of orientation. The y axis indicates the angle of anti-clockwise rotation about the central positive charge of the upper atom, while the x axis represents the offset towards the right-hand side of the diagram. Reproduced with permission from ref.[79]

It can be observed how the face-to-face geometry (with a 0° angle and 0 Å offset) causes repulsion. In the case of the edge-to-face conformation, where the rings are rotated by up to 90° relative to the other, and the offset stacked geometry, where they are laterally displaced from each other, π-σ attraction dominates, which means that they fall in the attractive band.

Studies on the effects of polarisation in these interactions have also been performed. In the case of stacked aromatic rings, if an electron-donating substituent (EDG) is present in one of the π-systems the electron density in the ring is increased, which in turn favours a stronger π-repulsion. In the case of an electron-withdrawing group (EWG), the electron density is removed from the ring, which decreases the π-repulsion between the systems. This favours a stronger interaction.[79]

In the case of edge-to-face interactions, the strongest attraction can be generated when an EWG is placed on the ring that contains the interacting proton (as it causes it to have a higher partially positive charge) and an EDG is present on the interacting π-system, which increases its basicity.[66]

π-π interactions play an important role in determining protein structure and stability. In β-hairpins, diagonal pairs (*i* to *j-2* in Figure 23) can interact through more than one geometry, while cross-strand pairs (*i* to *j* in Figure 23) interact through an edge-to-face conformation.[75]

The cross-strand edge-to-face interaction between a Phe-Phe pair contributes -2.3 kJ/mol to the ΔG.[66] Cross-strand Trp-Trp pairs, also in an edge-to-face geometry, provide the greatest stability to β-hairpins – this was used by Cochran *et al.* to design the Tryptophan Zipper peptide.[80]

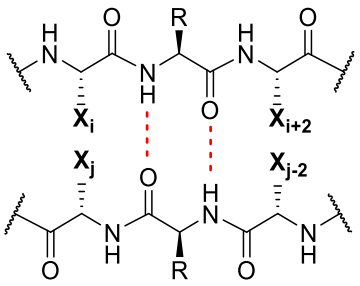


Figure 23. Cross-strand (*i*, *j*) and diagonal (*i*, *j-2*) interactions between side chains in a β-sheet.

1.2.6 Desolvation

Desolvation, or the removal of solvent from the protein structure, plays an important role in the folding process and also has an effect on the individual interactions that form the folded protein state. Theoretical studies have shown that there are desolvation barriers during the protein folding process (Figure 24) which are associated to the simultaneous expulsion of water from the protein core and the formation of new internal interactions. For proteins with 60-130 residues, the energetic effects from desolvation and solvation have been found to be in the order of 100 kJ/mol.[81]

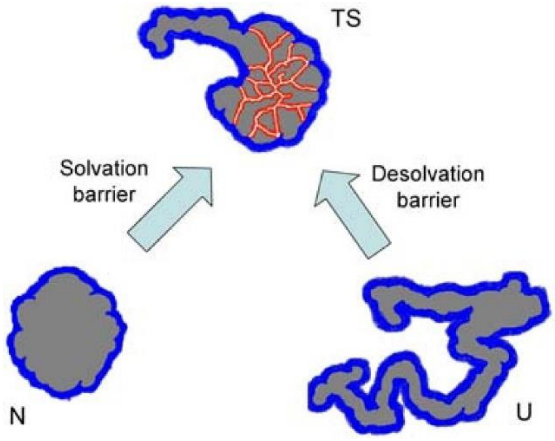


Figure 24. Solvation and desolvation barriers to protein folding and unfolding. The surfaces in blue in the native (N), unfolded (U), and transition (TS) states are exposed to the solvent, whereas the ones depicted in red represent broken internal contacts that are not solvated. Reproduced with permission from ref.[81]

One of the main driving factors of protein folding, the hydrophobic effect, takes place because of desolvation (discussed in Section 1.2.6.1). Cheung *et al*.[82] performed molecular dynamics simulations of the folding and unfolding of an SH3 protein and compared them to experimental values. They found that most of the folding process took place before water was removed from the hydrophobic core, and that a near-native intermediate was formed which contained a partially solvated core. The intermediate then underwent a cooperative desolvation step to remove the water and generate the native folded protein.

All non-covalent interactions in solution are influenced by desolvation. For example, this effect influences the energy of charge-charge interactions or salt bridges.[51] The free energy change upon formation of a salt bridge in a protein is made up of three components: $\Delta\Delta G_{bridge}$, the favourable contribution due to the formation of the salt bridge, $\Delta\Delta G_{protein}$, which arises from the electrostatic interactions of the salt bridge with the rest of the protein, and finally the $\Delta\Delta G_{desolvation}$, which is the unfavourable contribution from the desolvation of a buried salt bridge. It has been found that the $\Delta\Delta G_{desolvation}$ for a buried salt bridge is of 54 kJ/mol, whereas for an exposed salt bridge it is of 16 kJ/mol.[47] This highlights how the stability of a salt bridge is highly dependent on how buried it is within the protein interior.[2] For most salt bridges the desolvation cost is balanced by the electrostatic interactions between salt-bridging residues, but some have been found to be destabilising because of this effect.[83]

The formation of hydrogen bonds during folding is also connected to desolvation, as the hydrogen bonds between water and the backbone NH and CO groups as well as side chains have to be broken before new hydrogen bonds between residues can be formed.[84] Fernández *et al*.[85] studied the desolvation patterns of backbone hydrogen bonds in native proteins and found that as proteins fold, a simultaneous structuring and exclusion of water surrounding backbone hydrogen bonds takes place. This was proposed by the authors to compensate for the desolvation penalty of backbone polar groups. Although it has been proven that solvent-exposed hydrogen bonds stabilise folding, there is still controversy around whether buried ones are also favourable.[84]

The relationship between hydrogen bonding and desolvation was studied by Hunter,[86] who described the hydrogen bond interactions between a solute and a solvent as a function of two parameters denominated α and β. These parameters describe the hydrogen bond donating (α) and accepting (β) properties of functional groups on the solute as well as on the solvent (α_S

and β_S). When α and β are plotted, four quadrants are generated which describe the competition between solute–solute, solvent–solvent, and solute–solvent interactions (Figure 25). The red quadrants represent the case where hydrogen bonding interactions take place between solute and solvent, and therefore the interactions between solutes are unfavourable. In the blue quadrants solute-solute interactions predominate: the top right section corresponds to interactions between the most polar functional groups, while in the bottom left quadrant solute-solute interactions are favoured due to the solvophobic effect.

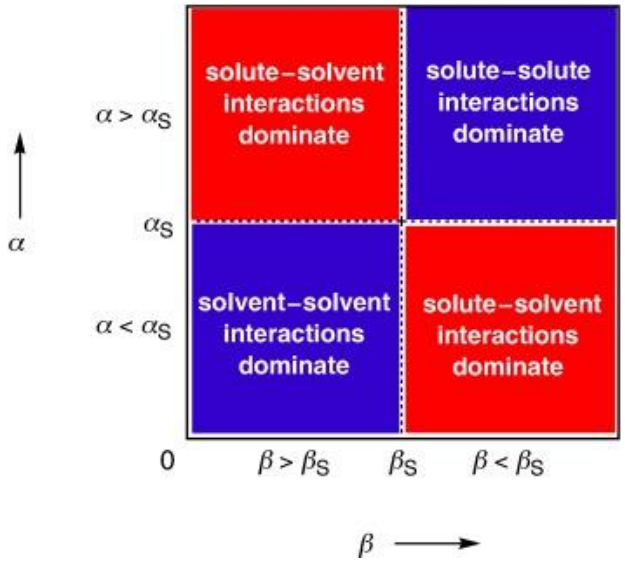


Figure 25. General profile for hydrogen bonding interactions between neutral functional groups in solution. Reproduced with permission from ref.[86]

The partitioning of the quadrants is different depending on the functional groups of the solute and the solvent. In the case of water, the values of α_S and β_S lie approximately in the middle of the scale, and therefore all quadrants are equal. This shows how both solute-driven and solvophobic interactions are important in this solvent. This model also describes the formation of a cage around non-polar solutes in water, which takes place to minimise the weak solvent–solute interactions in favour of the stronger solvent–solvent interactions.[86]

1.2.6.1 The hydrophobic effect

The hydrophobic effect is a special case of solvophobic effect that arises from the formation of clusters of hydrophobic residues in order to minimise their unfavourable contact with solvent water. Hydrophobic clusters can form in the core of proteins, where hydrophobic side chains are in close proximity and protected from the aqueous solvent, as well as on the protein surface.[87] In the latter case, the system rearranges for the non-polar side chains to be closer together, interact and reduce the area of the interface between them and water. A

driving force for the hydrophobic effect is the tendency for water molecules to hydrogen bond with each other, which takes place around non-polar residues and provides order to the system.[3]

1.3 Proteins and peptides

Proteins are one of the major components in all living systems, where they control biological activity. Peptides are relevant in a biological context, both in nature and in a design context, as they are being sought as therapeutics to solve issues encountered with resistance to small molecule drugs.

Proteins and peptides are polymers composed of amino acids linked by amide bonds. Each atom in a residue's backbone is labelled as C^{α} and H^{α} for the carbon and proton in between the NH and carbonyl groups, C^{β} and H^{β} for the following atoms in the side chain, etc (Figure 26).

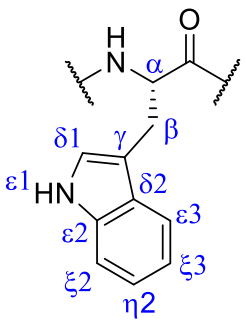


Figure 26. Designation of the side chain atoms in Trp.

The backbone of a polypeptide chain is described by the dihedral angles ϕ [C(=O)-N-C^{α}-C(=O)], ψ [N-C^{α}-C(=O)-N] and ω [C^{α}-C(=O)-N-C^{α}] (Figure 6, Figure 27). The accessible regions for the dihedral angles ϕ and ψ are rather restricted due to the steric interactions that take place during rotation. These regions were described by Ramachandran[88] and they are usually represented through a ϕψ-map or Ramachandran plot (Figure 27). The two most populated regions correspond to the right-handed α-helix (near -60°, -60°) and the β-sheet (near -90°, +120°).

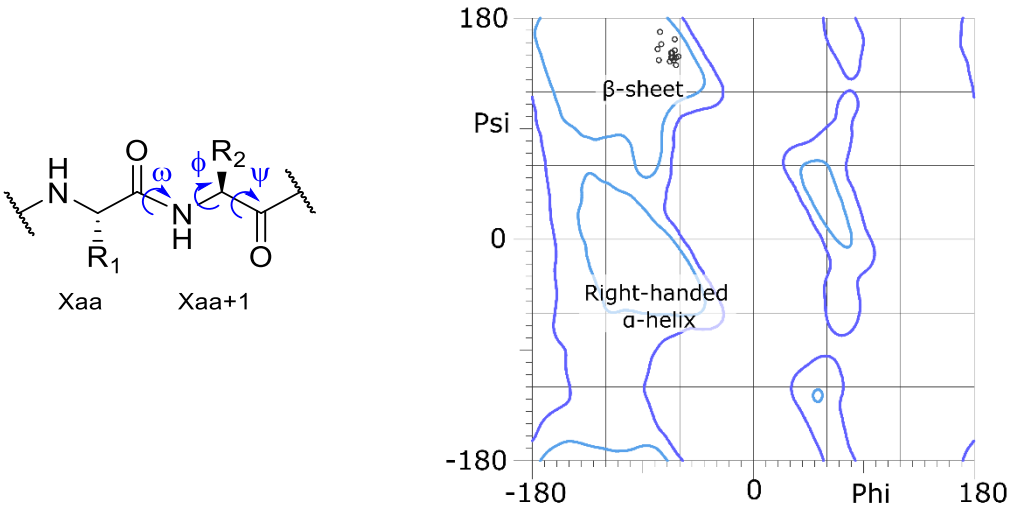


Figure 27. Left: dihedral angles ω, ϕ and ψ in a peptide. Right: Ramachandran plot for collagen (PDB accession code 1BKV), generated using MolProbity.[89,90]

The amide bond's partial double bond character restricts the available energy minima for the ω angle to two, therefore it can only adopt *cis* (ω = 0°) and *trans* (ω = 180°) conformations (Figure 28; this nomenclature will be used in the context of the peptide bond hereinafter). These two states are separated by a rotational barrier corresponding to a high energy state of ω = 90°.[91]

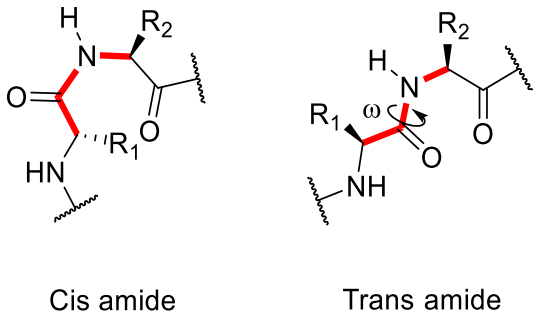


Figure 28. *Cis* and *trans* isomers of a secondary amide peptide bond.

In secondary amide bonds the *trans* configuration is the most favoured because repulsion between the C^{α} atoms is minimised. This causes the energetic difference between the two conformers (ΔG) to be approximately 10 kJ/mol, with a rotational barrier (ΔG_A) of 84 kJ/mol (Figure 29).[5] Of all canonical amino acids, proline is the only one that establishes a tertiary amide bond, the prolyl bond. In this case both conformers meet a comparable steric hindrance, which causes the energy difference between them to be smaller, approximately 2 kJ/mol (with a rotational barrier of 54 kJ/mol).[5,92]

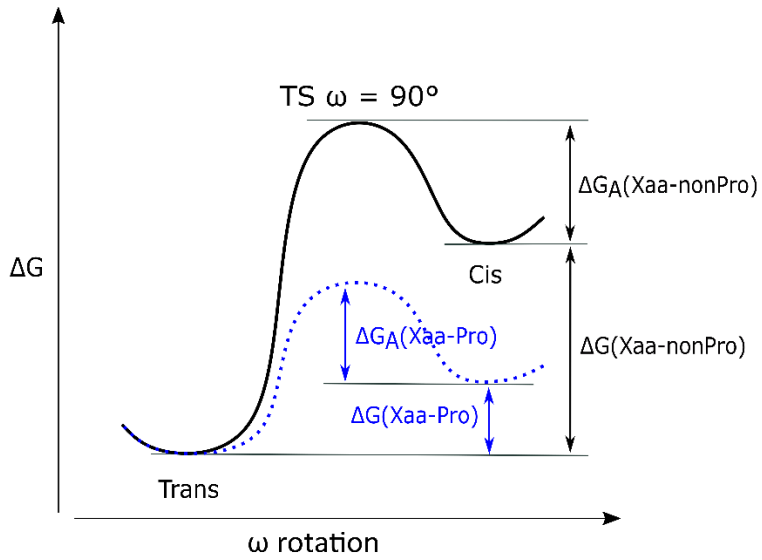


Figure 29. Energy diagram for the isomerisation in Xaa-Pro (blue) and Xaa-nonPro peptides (black). The energy barrier (ΔG_A) as well as the energetic difference between the *cis* and *trans* states (ΔG) is smaller for Xaa-Pro peptides.

1.3.1 Secondary structure

Proteins are arranged in space by forming secondary, tertiary and quaternary structures. The secondary structure is the local conformation of the chain and the two latter correspond to the overall folding of the polypeptide and the formation of subunits, respectively.

There are three main secondary structures in proteins: the α-helix, β-sheet and reverse turns. In an α-helix residues adopt set values for their dihedral angles, which induces a pitch along the helix (vertical distance between consecutive turns, Figure 30). This means that the residues are offset from each other and 3.6 residues are found per turn in a regular right-handed helix. The structure is hydrogen bonded and these interactions take place between the carbonyl oxygen of residue *i* and the NH of residue *i+4*. α-Helices compose over 30% of the protein secondary structures present in nature:[93] for example, they form coiled-coils in keratins and triple helices in collagen.[3] They are also involved in protein-protein interactions and have therefore been an attractive target for the development of mimetics[94] of this secondary structure.

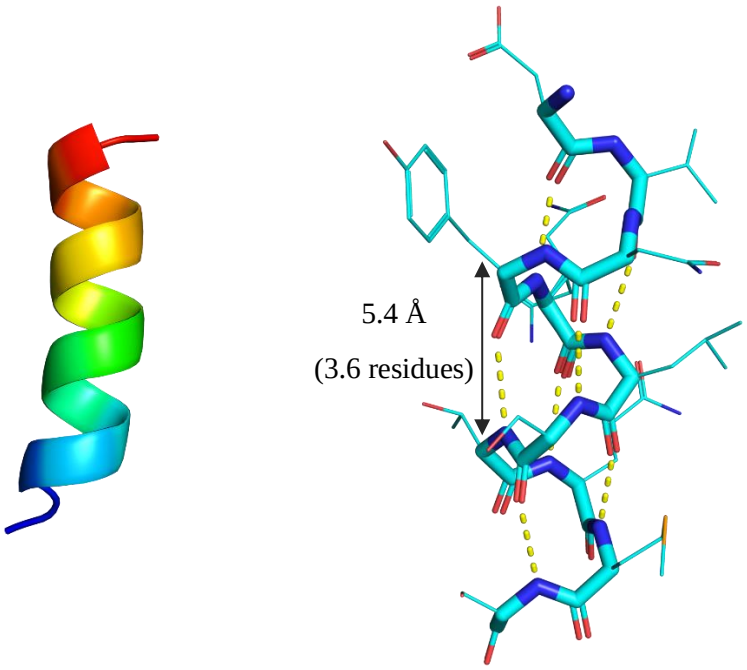


Figure 30. Left: general structure of an α-helix (PDB accession code 111L). Right: α-helix showing the hydrogen bonding pattern in yellow and the distance between atoms (PDB accession code 3IEE).

The β-strand is a more extended conformation with a smaller pitch than helices (0.7 nm). The lack of stabilising interactions within the structure causes a single β-strand to not be stable, the β-sheet being a more commonly observed motif. β-Sheets are formed by two or more β-strands that are hydrogen bonded to each other (Figure 31), which confers much more stability to the structure. The orientation of the strands can be parallel, antiparallel or mixed. These structures play key structural roles in fibrous proteins, like silk,[95,96] and they also participate in DNA-protein interactions[97] (for example gene regulation processes[98]) and protein-protein interactions (PPIs).[99]

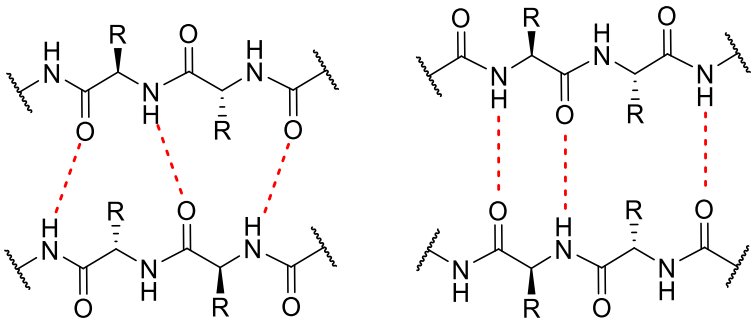


Figure 31. Hydrogen bonding pattern in parallel (left) and antiparallel (right) β-sheets.

Parallel β-sheets are formed by β-strands parallel to each other, with the termini pointing in the same direction, and antiparallel sheets contain β-strands in opposite directions (Figure 32). The two dispositions differ in the backbone dihedral angles adopted as well as the

hydrogen bonded pattern between strands. Completely antiparallel sheets are more commonly observed than parallel ones, as the latter are only formed if four or more strands are present. Conversely, antiparallel β-sheets formed by only two strands are often observed – the most simple motif is the β-hairpin, where two antiparallel β-strands are joined by a reverse turn (Figure 32).

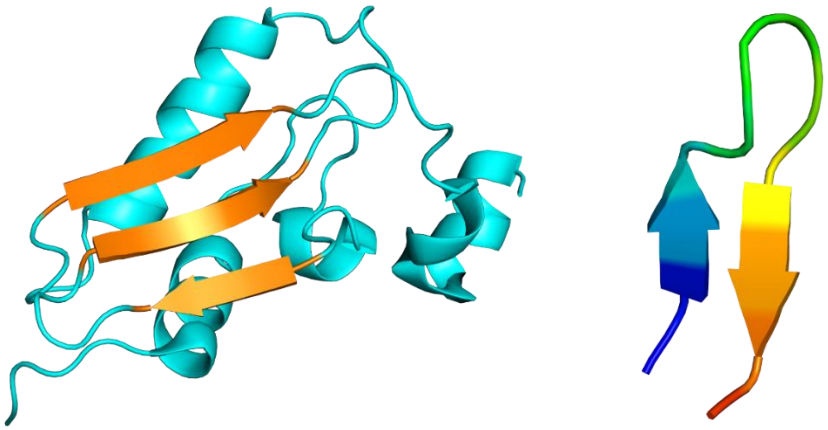
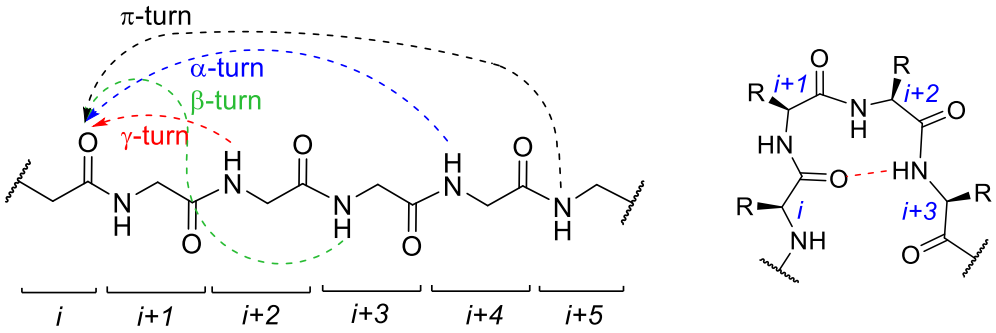


Figure 32. Parallel and antiparallel β-sheets (left, PDB accession code 1WCJ), and β-hairpin (right, PDB accession code 1LE0).

Reverse turns are key elements of secondary structure because they allow the polypeptide chain to change direction. They are usually found on the surface of proteins[100,101] and are therefore implicated in molecular recognition processes.[100,102] Turns also nucleate folding in β-hairpins[103–105] and can be involved in the rate-determining step of folding in larger proteins.[106] The stability of these motifs greatly affects the stability of the ''larger'' protein conformation.[107,108] The different types of turn are classified according to the number of residues present and which ones are hydrogen bonded, as well as the dihedral angles ϕ and ψ of the central residues.

There are many kinds of reverse turns: the smallest are δ-turns, which are composed of two residues hydrogen bonded from NH(*i*) to CO(*i+1*) and γ-turns, which involve three residues with a hydrogen bond between CO(*i*) and NH(*i+2*). β-Turns involve four residues and are usually hydrogen bonded from CO(*i*) to NH(*i+3*), although 25% of them have been found to not contain the hydrogen bond (Figure 33).[109,110] A more general definition for β-turns[111] uses the distance between $C^{\alpha}(i)$ and $C^{\alpha}(i+3)$, which is always smaller than 7 Å. α-Turns are formed by five residues and have a distance smaller than 7 Å between $C^{\alpha}(i)$ and $C^{\alpha}(i+4)$. Like in β-turns, the hydrogen bond between *i* and *i+4* is not always present. Lastly, the loosest reverse turns are π-turns, which are composed of six residues and contain a hydrogen bond between *i* and *i+5*.[112]



π-turn: 16-membered ring hydrogen bond: NH (*i*+*5*) → CO (*i*)

α-turn: 13-membered ring hydrogen bond: NH (*i*+*4*) → CO (*i*)

β-turn: 10-membered ring hydrogen bond: NH (*i*+*3*) → CO (*i*)

γ-turn: 7-membered ring hydrogen bond: NH (*i*+*2*) → CO (*i*)

Figure 33. Types of reverse turns (left) and representation of a β-turn (right).

1.4 β-Turns

β-Turns are secondary structure elements that were first recognised and described by Venkatachalam in 1968.[113] Out of all turn types, the β-turn is the most thermodynamically favourable and most common in proteins. They are key motifs in protein folding[108,114] because they cause the polypeptide chain to fold back on itself by nearly 180° (Figure 34), which provides a change in direction. Their role as nucleation sites is observed in β-hairpins, where a β-turn connects the two strands of an antiparallel β-sheet and promotes folding. They are also important elements in protein function as they are frequent sites for molecular recognition.[100,101]

These turns comprise four consecutive residues, *i* to *i*+*3*, and are generally defined by a hydrogen bond being present between the carbonyl of residue *i* and the amide NH of residue *i*+*3* (Figure 34).[110,111] 25% of β-turns are open,[109] in which case the distance between the $C^{\alpha}(i)$ and $C^{\alpha}(i+3)$ is always less than 7 Å. These turns are commonly found connecting other secondary structure elements like β-hairpins.

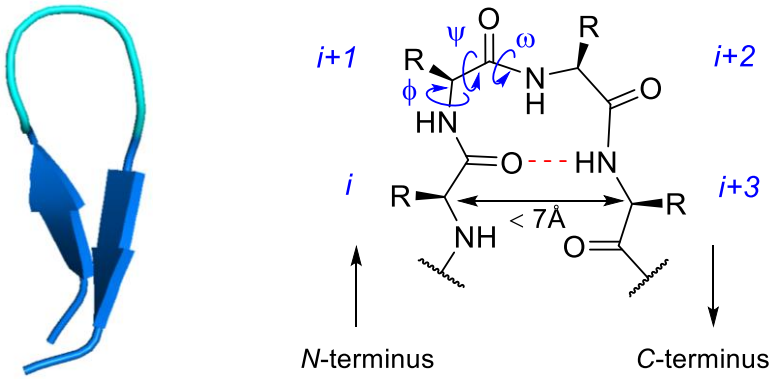


Figure 34. Left: β-hairpin containing a β-turn, showing the chain reversal. PDB code 1LE0. Right: general scheme of a β-turn, showing the dihedral angles (in blue, for further definition of the dihedral angles see Figure 6), the residues *i* to *i+3* and indicating the hydrogen bond from *i* to *i+3* (red).

There are many types of β-turns, which are defined by the dihedral angles phi ϕ and psi ψ of the backbone residues *i+1* and *i+2*. β-Turn types were originally described by Venkatachalam[113] and their classification has since then evolved to include 12 different turns[109,110] – the mean dihedral angles corresponding to all types are presented in Table 1.

Table 1. Turn types with the ϕ and Ψ dihedral angles for the *i+1* and *i+2* residues.[115]

Turn type	ϕ_{i+1}	ψ_{i+1}	ϕ_{i+2}	ψ_{i+2}
I	-60°	-30°	-90°	0°
I'	+60°	+30°	+90°	0°
II	-60°	+120°	+90°	0°
II'	+60°	-120°	-90°	0°
III	-60°	-30°	-60°	-30°
III'	+60°	+30°	+60°	+30°
V	-80°	+80°	+80°	-80°
V'	+80°	-80°	-80°	+80°
VIa1	-60°	+120°	-90°	0°
VIa2	-120°	+120°	-60°	0°

Table 1. Turn types with the ϕ and Ψ dihedral angles for the *i+1* and *i+2* residues.[115]

Turn type	ϕ_{i+1}	ψ_{i+1}	ϕ_{i+2}	ψ_{i+2}
VIb	-135°	+135°	-75°	+160°
VIII	-60°	-30°	-120°	+120°

Type VI β-turns stand out from the other types because the residue *i+2* is always a *cis* proline (Figure 35) and therefore the angle ω(*i+1*) is 0° instead of 180°. The presence of a *cis* peptide bond is unique in this kind of turn and it is a relevant motif in protein folding and function:[116–118] they have been found to mediate protein-protein interactions[119,120] and are recognition elements for protein receptors.[100]

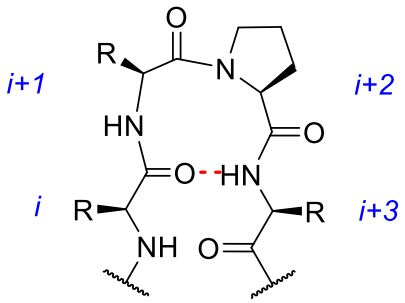


Figure 35. Type VI β-turn containing a *cis* Pro at the *i+2* position.

1.5 β-Hairpins

β-Hairpins are the simplest antiparallel β-sheet and one of the most basic structural motifs adopted by peptides in aqueous solution. These supersecondary structure motifs are composed of two antiparallel β-strands that are hydrogen-bonded throughout the backbone and connected by a loop region (Figure 36). In most systems this region adopts β-turn topologies.

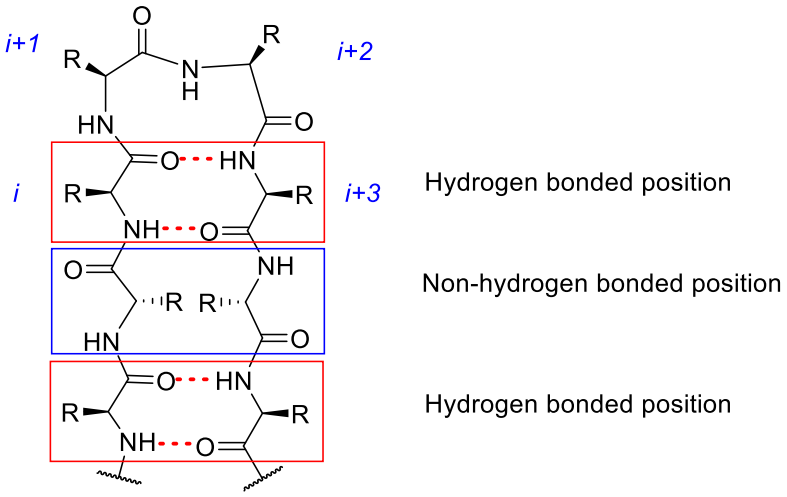


Figure 36. General structure of a β-hairpin. The hydrogen bonds are represented in red, and the hydrogen-bonded positions are in red boxes. The non-hydrogen bonded positions are shown in blue boxes. The β-turn residues are labelled in blue.

Although types I and II are the most common β-turn conformations, they are rarely observed in β-hairpins. This is due to the incompatibility between the left-handed twist in the turn and the right-handed twist that takes place in β-hairpins.[74] Instead, the mirror image turn types I' and II' have dihedral angles that are complementary to the twist in β-hairpins, thus they are almost exclusively found in these systems.[121] Gellman *et al*.[122] proved this in peptides of the sequence MQIFVKSxxKTITLKV-NH_2, where xx was (L-P)-(L-A), (D-P)-(L-A), (D-P)-(D-A), (D-P)-G, and (L-P)-G. They showed that replacing L-Pro with D-Pro changed the turn type from I or II to the opposite I' or II', which made the turn compatible with the hairpin twist and therefore stabilised the system.

Turn sequence also has a strong influence in defining hairpin conformation: it has been shown that using proteinogenic amino acids, Asn-Gly turns are one of the most effective at inducing β-hairpin formation,[121] as they favour type I' β-turns. The (D-Pro)-Gly motif is also widely used to nucleate β-hairpin formation, and it has been shown to be a stronger promoter of folding,[123] forming better-defined structures than Asn-Gly turns for smaller β-hairpin peptides.

The stability of β-hairpins is attributed to the interplay of both interstrand hydrogen bonding and hydrophobic interactions. These two components were used by Cochran and co-workers to plan one of the most successful β-hairpin designs to date, the tryptophan zipper[80] (Figure 37). This motif is based on a pattern of polar residues in the hydrogen-bonded positions and tryptophan residues in the non-hydrogen bonded positions. This causes the formation of two

faces in the system, including a cross-strand tryptophan stacking pattern, which confers substantial stability to the system.

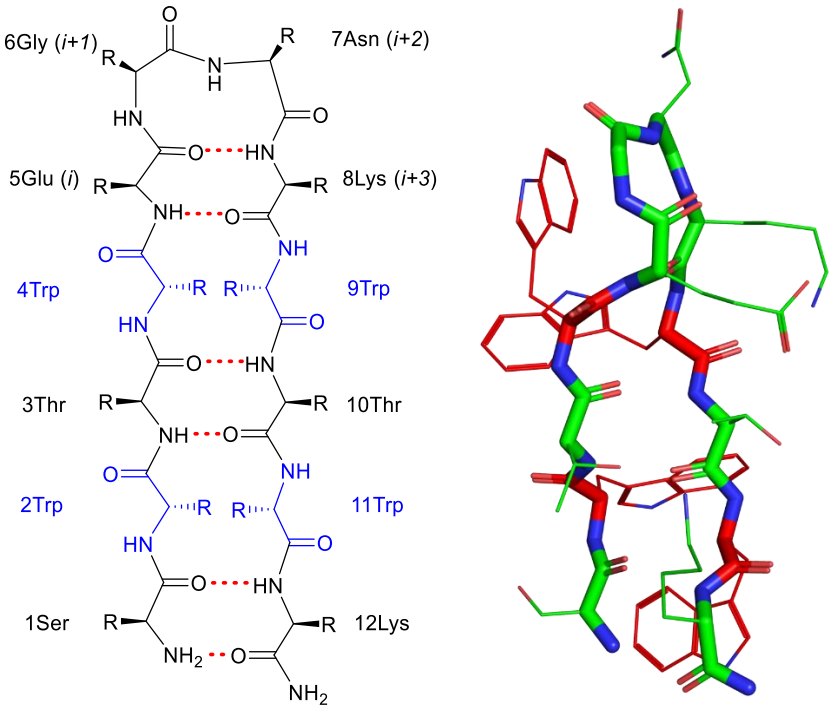


Figure 37. Left: TrpZip1 structure showing the hydrogen bonded positions in red. Right: NMR structure showing the Trp residues in red and hydrophilic residues in green.

Various studies have been performed to elucidate the impact of hydrophobic interactions in the stability of hairpin systems,[124,125] which have shown the key role played by the coupling of cross-strand hydrophobic residues. It has been observed that Trp-Trp interactions have a bigger impact than interactions between other residues, and that they are additive and have a geometrical aspect.

The effect of disulfide bridges on hairpin stability has also been studied – Santiveri *et al.*[126] found that their contribution to stability is dependent on their placement in the β-hairpin. If the bridged cysteines are placed at non-hydrogen bonding positions, the disulfide favours folding and it can contribute to the system's conformational stability by 4-5 kJ/mol. Its closeness to the turn region will modulate the stabilising effect. Conversely, forming a disulfide bridge between Cys residues at hydrogen-bonding positions between antiparallel strands is not geometrically favourable, which causes strain in the system. This causes the disulfide's contribution to stability to be small or even destabilising.

β-Hairpins are very interesting structures as they have varied roles in biological processes: they are involved in molecular recognition events,[127] protein-protein interactions[128,129] and protein-DNA interactions,[130] and have been found key in some T-cell activation processes.[131]

1.6 Designed β-hairpins

Tools to induce and stabilise β-turns are being sought for as peptides with this motif, especially β-hairpins, have a wide range of applications, from the synthesis of new therapeutic agents[132,133] to using them as scaffolds in the design of switchable peptides.[134,135] The design of β-turn and β-hairpin mimics is not unique, as the mimicry of other elements of protein secondary structure has also attracted attention.[136] One of the most explored and significative is the mimicry of α-helices.[94] Completely unnatural α-helix backbones have been designed,[137,138] and stapling[139] and hydrogen bond surrogate[140] strategies are commonly used to stabilise the helical backbone.

β-Hairpin and β-turn stability can be improved through various strategies:[141] using a hydrogen bond or peptide bond surrogate, employing macrocyclisation or introducing a non-natural turn moiety.

1.6.1 Hydrogen bond surrogate approach

The hydrogen bond surrogate (HBS) approach was originally developed to stabilise α-helices,[142] and has been widely used in this field. In this strategy, a main chain hydrogen bond is replaced by a covalent linker that can be of various kinds, from olefins, disulfide bridges, to functionalities accessed through ligation chemistry like hydrazones or oximes.

Arora *et al.*[143] were the first to apply this strategy to β-hairpin backbones, where they replaced the *N*- to *C*-terminus cross-strand hydrogen bond with a covalent HBS linker. Hydrocarbon, thioether and disulfide surrogates (Figure 38) were studied, to find that the two latter give better hairpin stabilities. The biggest benefit of this strategy is that all side chains along the β-sheet are unmodified, as only one main chain hydrogen bond is replaced with a covalent linkage at the termini.

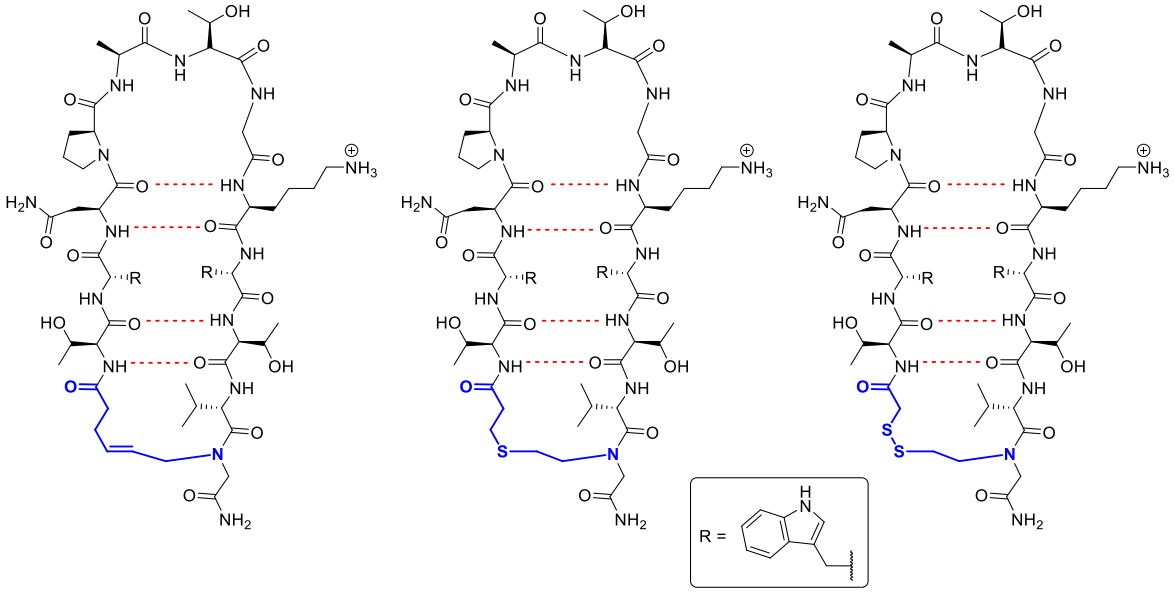


Figure 38. β-Hairpins designed by Arora *et al.*[143] containing a hydrocarbon (left), thioether (middle) and disulfide (right) HBS linkage, in blue. The hydrogen bonds across the hairpins are represented in red.

1.6.2 Peptide bond surrogates

A wide variety of systems can act as peptide bond surrogates[144] – these range from thioamides, esters and alkenes to triazoles[145] (Figure 39) as well as Schiff bases.

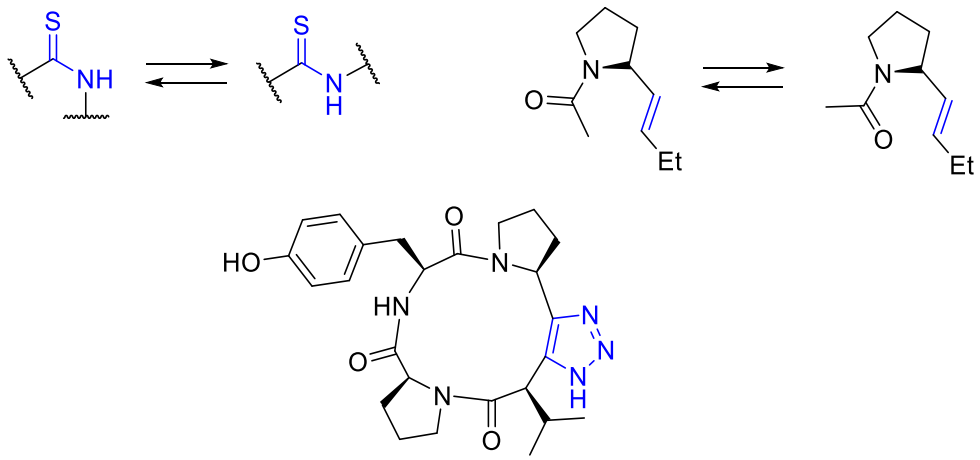
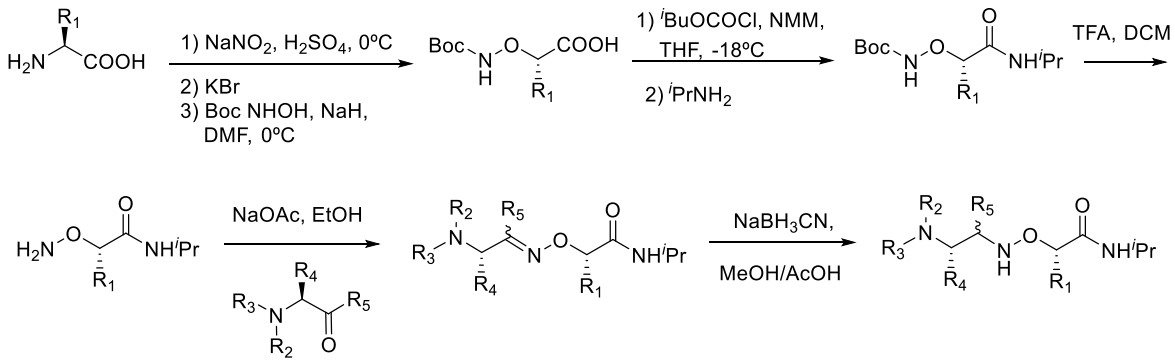


Figure 39. Thioamide, alkene and triazole-based peptide bond surrogates.

Vanderesse *et al.*[146] used hydrazide, oxime, amidoxy and *N*-alkoxy amine moieties as surrogates to induce a γ-turn. These functionalities were introduced as amide bond surrogates between positions Xaa and Yaa of dipeptides of the formula RCO-Xaa-Yaa-NHR'. The authors observed that the oxime-linked sequences could form an extended conformation, arising from the *E*-stereoisomer, or a β-like conformation, adopted by the *Z*-stereoisomer. The *N*-alkoxy amine dipeptides were found to be quite flexible, existing in equilibrium

between an open and a folded conformation. Finally, they found that the amidoxy and hydrazide linkages were very similar amide surrogates, as they stabilised analogous γ-like folded structures.

These dipeptides were accessed through the synthesis of the *N*-alkoxy amine-containing unnatural amino acid, followed by a ligation and reductive amination to generate the oxime- and *N*-alkoxy amine-containing peptides, respectively (Scheme 3).



Scheme 3. Synthesis of the oxime and *N*-alkoxy amine peptides.[146]

Longer hexapeptides were also studied (Figure 40), containing the amidoxy, oxime and *N*-alkoxy amine linkers, which were synthesised through fragment condensation between the *N*-alkoxy amine-containing peptide and the aldehyde-containing fragment. These peptides were evaluated as potential activity inhibitors of human leukocyte elastase.

Ala-Ala-Pro-*link*-Val-Ala-Ala-NH^{i}Pr

Ala-Ala-Pro-Val-*link*-Ala-Ala-NH^{i}Pr

Figure 40. Hexapeptide sequences studied by Vanderesse *et al*.[146]

1.6.3 Macrocyclisation

The stabilisation of β-hairpins can be achieved through covalent crosslinks that do not replace a hydrogen bond. The crosslink can be introduced as a head to tail macrocyclisation as well as a local staple. Local aliphatic staples (Figure 41), for example, have been used to stabilise a β-turn[147] and to promote a *cis* configuration in tripeptides of the sequences Pro-Leu-Pro-NH_2 and Pro-Pro-Pro-NH_2.[148]

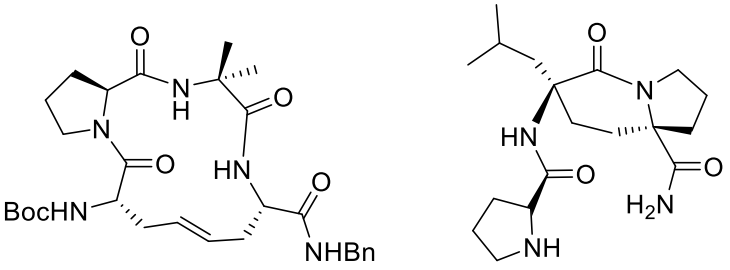


Figure 41. β-Turns stabilised with aliphatic staples by Miller and Grubbs[147] (left) and Vartak *et al*.[148] (right).

Some of the most useful crosslinks are introduced between residue side chains along the β-hairpin strands. Some of the crosslinks, like disulfide bridges, require to be placed at non-hydrogen bonding positions in the strands in order to be stabilising, whereas others, like triazoles, are more flexible in their placement.[149] Disulfide bridges can be introduced between two cross-strand Cys residues at any non-hydrogen bonding position along the hairpin, although they are most stabilising when cyclising between the *C*- and *N*-termini.[150,151]

1,4-Disubstituted 1,2,3-triazoles have also been widely used to stabilise β-hairpins, both between the *C*- and *N*-termini[152] and at other positions along the strands (Figure 42).[153,154] These linkages are accessed through a copper(I)-catalysed azide-alkyne cycloaddition between two amino acids containing an azide and an alkyne moiety. Usually this reaction is performed in solution, but Holland-Nell and Meldal have shown how two triazole bridges can be introduced on solid support.[155]

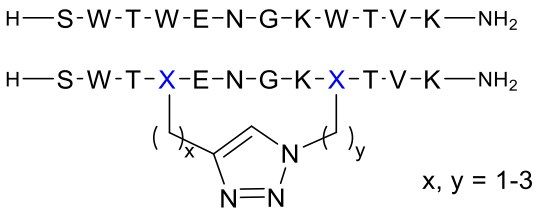


Figure 42. Triazole-bridged β-hairpin studied by Celentano *et al*.[154]

1.6.4 β-Turn mimics

The development of β-turn mimics is one of the most explored strategies to control folding and provide a higher stability to β-hairpins. Many turn mimics have been developed along the years, from the initial use of D-amino acids to the introduction of partial and completely unnatural fragments in the turn region.[156–158] This strategy provides a very useful tool for the study of protein conformation as well as the development of potential therapeutic agents,[141] to develop new antimicrobial peptides[159] as well as protein-protein interaction inhibitors.[160]

The dipeptides (D-Pro)-(L-Pro)[161] and (D-Pro)-Gly[122] are two of the most widely used β-turn mimics[162] due to their straightforward introduction in the peptide sequence. They stem from the Pro-Xaa bond (where Xaa = Pro, Gly or Asn), which is a sequence commonly found in β-turns in natural proteins. These dipeptides contain a D-Pro instead of L-Pro at the *i+1* position, which changes the turn conformation from a type I or II to a I' or II' β-turn. This makes the turn more compatible with the right-handed twist in a β-hairpin and provides more stability.

Other turn mimetics comparable to the (D-Pro)-Xaa dipeptides in β-hairpin stabilisation are δ-ornithine[163] and the (D-Phe)-2-Abz motif (Figure 43),[164] which fuses the ring from 2-amino benzoic acid into the backbone.

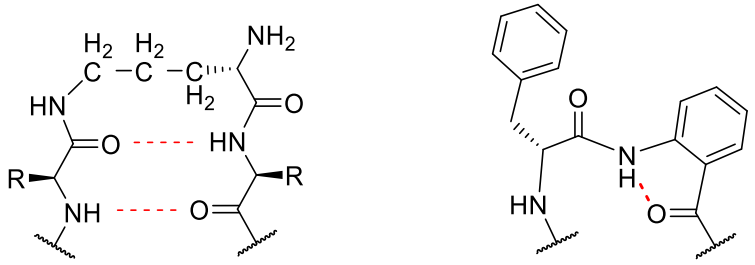


Figure 43. Left: δ-ornithine β-turn. Right: (D-Phe)-2-Abz turn. The hydrogen-bonded positions are shown in red.

δ-Ornithine has also been employed in combination with Hao, a tripeptide β-strand mimic, to generate stable, folded cyclic β-sheets in aqueous medium (Figure 44).[165] The two δ-ornithine moieties at each end of the hairpin form β-turns and nucleate folding, whilst Hao has a planar geometry that orients the hydrogen bond donors and acceptors in a favourable way.

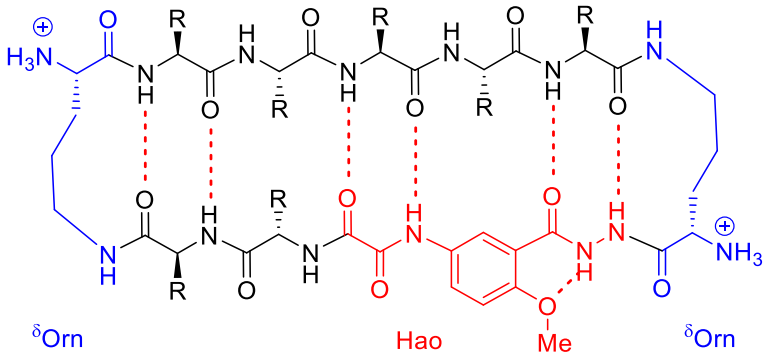


Figure 44. β-sheet containing δ-ornithine and Hao, studied by Woods *et al*.[165]

The Aib motif (α-aminoisobutyric acid) has also been widely used to nucleate β-hairpin formation. It was developed by the Balaram group, who showed it forms a type I' β-turn when in combination with D-Ala in the solid state.[166] Since that first finding, the Aib-(D-Ala) motif has been shown to form β-turns in water[167] and Aib has been used as a β-turn inducer in combination with other amino acids like D-Pro as well as Gly.[168]

Mimics of type VI β-turns have also been developed using a 1,5-disubstituted[1,2,3]triazole,[145] which eliminates the possibility of *cis/trans* isomerisation and acts as a surrogate for *cis* Pro. 1,4-Disubstituted[1,2,3]triazoles have also been used in tetrapeptides, where various linker lengths have been explored to find that three methylene units provide the most stable β-turn (Figure 45).[169]

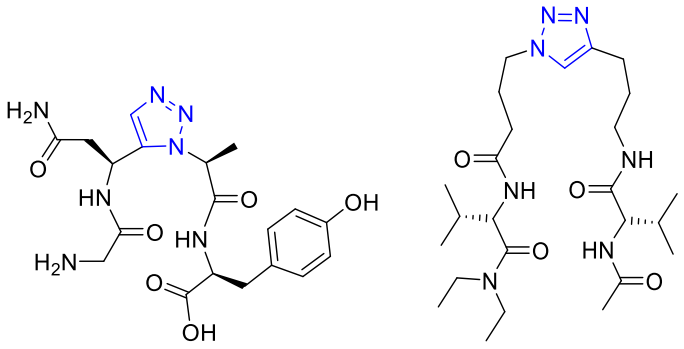
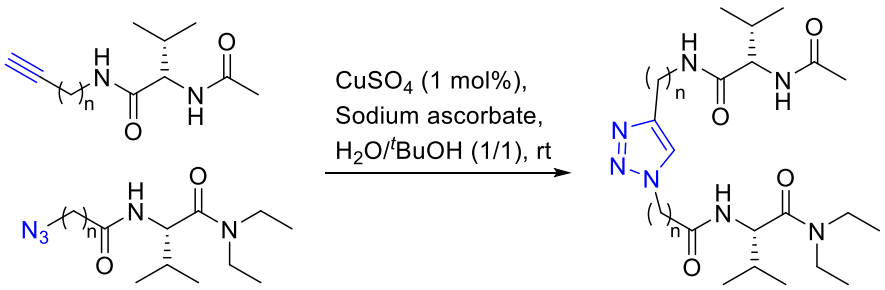


Figure 45. Left: Asn-1,5-triazole-Ala as a *cis* Pro surrogate.[145] **Right: 1,4-triazole with a 3-carbon linker.**[169]

The 1,4- disubstituted[1,2,3]triazole systems are accessed by means of a copper(I)-catalysed cycloaddition (Scheme 4):



Scheme 4. Synthesis of 1,4-triazole β-turn mimics,[169] where n = 1-4.

Peptidomimetics based on a semicarbazide (Aza) as well as a α-amino-γ-lactam moiety (AgI and HgI) have been developed (Figure 46), which inspired the design of the Aid and Nai residues by Lubell *et al.*[170] Tetrapeptides synthesised with the Aid residue form compact β-turns with geometries of both type II and II' being observed.

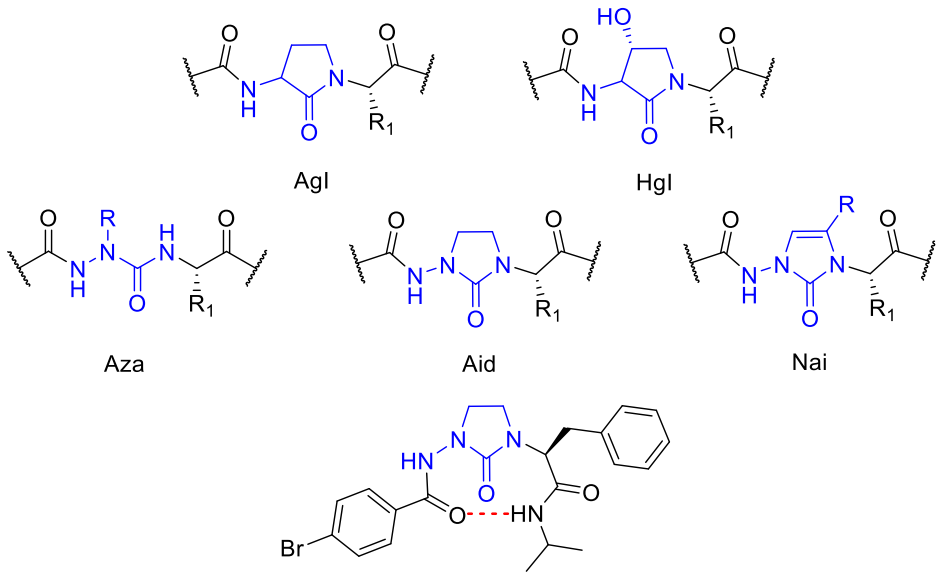
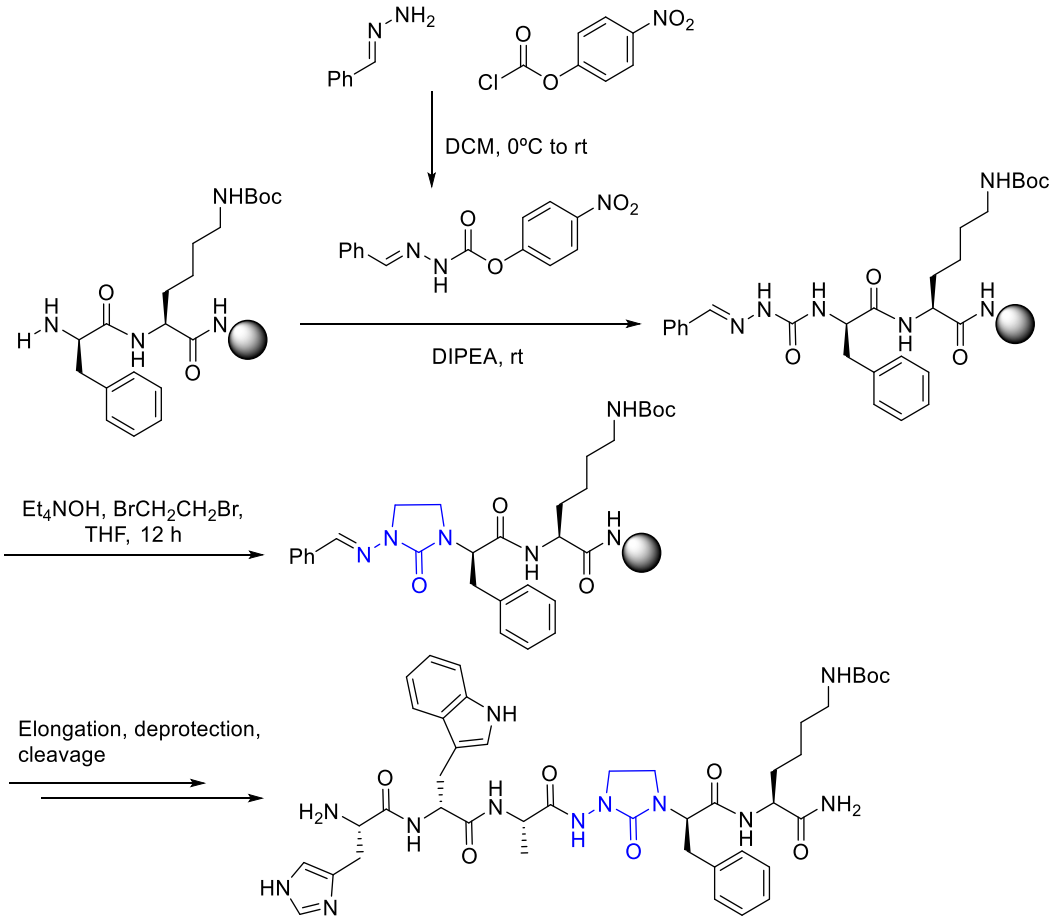


Figure 46. Semicarbazide and α-amino-γ-lactam-based peptidomimetics. Bottom: Aid-containing peptide studied by Lubell *et al.*,[170] showing the β-turn hydrogen bond in red.

Longer peptides containing the Aid residue can be synthesised *via* solid-phase peptide synthesis (SPPS, Scheme 5).[171] This residue is accessed through the coupling of a hydrazone onto the peptide sequence, followed by a bisalkylation, deprotection and standard solid phase peptide synthesis.



Scheme 5. Synthesis of Aid-containing peptides.[171]

Fink *et al.*[172] have designed a type I β-turn mimic through covalently attaching the positions *i+1* and *i+2* in the turn (Figure 47) – this constrains the ϕ and ψ angles to within 30° of the ideal angles for this type of turn. The mimic has been evaluated within tetrapeptides, which are synthesised by means of a ring closing metathesis to generate the turn unit, followed by solution-phase peptide synthesis.

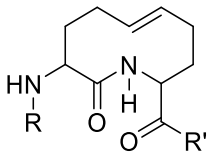


Figure 47. β-Turn mimic designed by Fink *et al.*[172]

Bicyclic systems that replace the *i+1* and *i+2* residues in a β-turn are a common way of designing β-turn mimics. One of the first bicyclic β-turn mimics was designed by Nagai and Sato[173] and it has recently been improved upon by Eckhardt *et al.*,[174] who introduced two hydroxyl groups into the mimic (named Hot=tap) to enhance its turn-inducing properties (Figure 48).

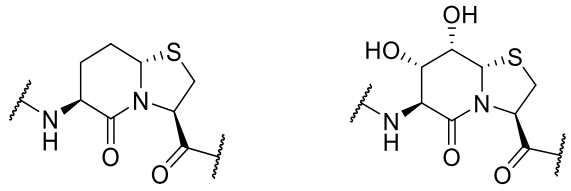
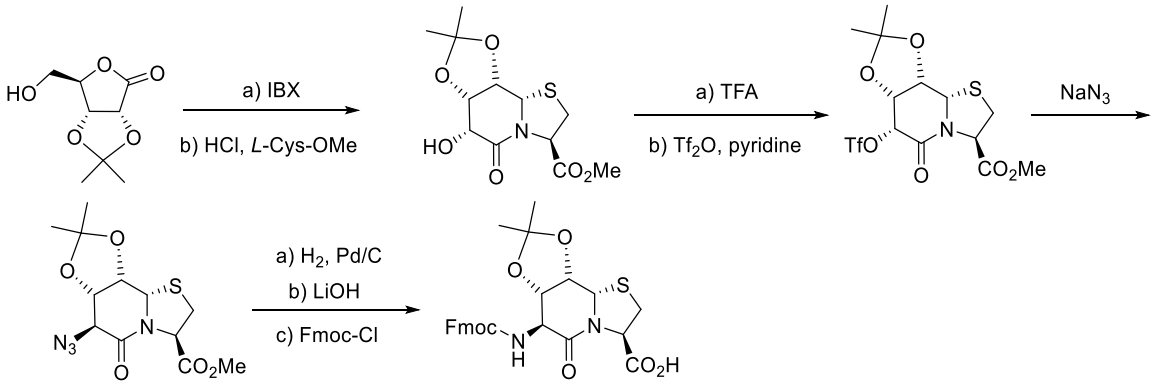


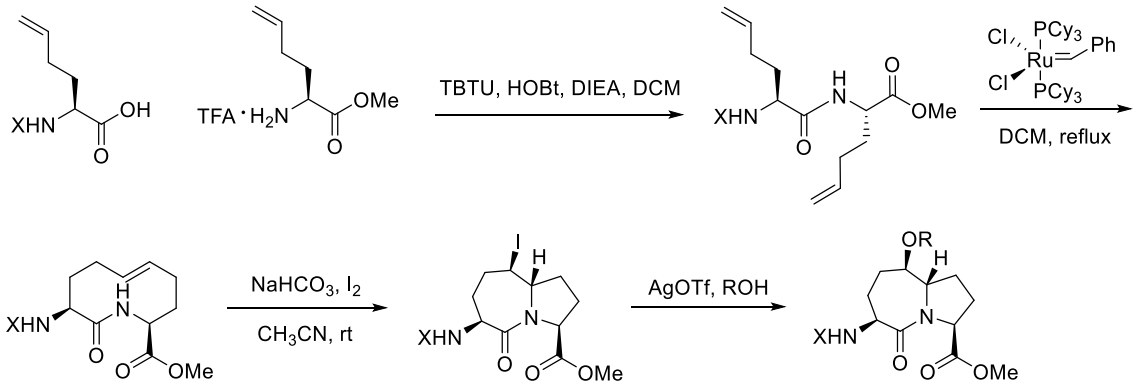
Figure 48. β-Turn mimics designed by Nagai *et al*.[173] (left) and Eckhardt *et al*.[174] (right).

Hot=tap has been used in the two turn sites within cyclic hexapeptides as well as in various sites within a fibritin-foldon miniprotein, in both cases showing successful mimicry of the β-turns. This mimic is introduced in the peptide sequence through standard SPPS, which requires the synthesis of the unnatural amino acid (Scheme 6).



Scheme 6. Synthesis of Hot=tap.

Other bicyclic systems are in the literature, like the dipeptide designed by Godina *et al.*[175] based on an azabicyclo[5.3.0]alkanone skeleton (Scheme 7), but most of these systems involve a multi-step synthesis, which limits their utility in peptide and protein science.



Scheme 7. Synthesis of the general scaffold of the β-turn mimic designed by Godina *et al*.,[175] where X = Fmoc or Boc and R = Me, Et, ^{i}Pr, Ph or H.

All in all, the development of β-turn mimics is a widely explored field, with numerous and diverse strategies having been developed previously. Many of them require multi-step

synthetic routes to synthesise the unnatural unit or to install it on a peptide backbone, thus the most commonly used motifs in the literature are D-Pro in combination with L-Pro, Gly or Asn as well as the unnatural amino acid ornithine.

1.7 Applications of β-turn mimics

β-Hairpin peptides that contain unnatural β-turn mimics have a wide range of applications, from their use in conformational studies to utilising them as therapeutic agents.[156] These units provide control over the conformation of a peptide and facilitate the stabilisation of structures that would otherwise not have a defined conformation. The ability to synthesise systems that are smaller than proteins but still fold in a defined manner makes β-turn mimics a useful tool to model aspects of protein structure. Selected examples of the applicability of these mimics are presented in this section.

1.7.1 β-Turn mimics to understand protein folding

The effect that using different residues in a β-turn can have in the stability of the β-hairpin has been widely explored in the literature, and it has been shown how some amino acids favour the formation of a turn more than others. Cochran *et al.*[80] studied the thermal stabilities of designed tryptophan zipper peptides (TrpZip, Table 2) and found that the melting temperatures increased when using an Asn-Gly turn in comparison to a Gly-Asn turn. The thermal stability increased even further when a D-Pro was introduced at the *i+1* position of the turn, generating a (D-Pro)-Asn turn.

Table 2. Tryptophan zipper peptides studied by Cochran *et al.*[80] and their melting temperatures.

	Peptide sequence	Melting temperature (K)
TrpZip1	SWTWEGNKWTWK	323.0 ± 0.3
TrpZip2	SWTWENGKWTWK	345.0 ± 0.1
TrpZip3	SWTWE(D-Pro)NKWTWK	351.8 ± 0.2

Building on the knowledge that different amino acids affect the stability of β-hairpin systems, unnatural β-turn units can be introduced to understand how folding takes place. Nowick *et al.*[176] used homodimers of peptides containing the Orn(^{i}PrCO-Hao) unit (Figure

49) to measure the interaction differences between various amino acid side chains, in order to understand how molecular recognition takes place between β-sheets. They found that when a Thr-Thr and a Val-Val pair were established, the stability of the homodimers was 2.5 kJ/mol higher than when two Thr-Val pairs took place. In addition, when a Ser-Ser pair and a Val-Val pair were established, the stability increased by 1.7 kJ/mol compared to when two Ser-Val pairs took place. This demonstrated how pairing in β-sheets is sequence selective, and suggested that hydrogen bonding with Thr and Ser side chains might cause the selectivity in the dimerisation of these peptides.

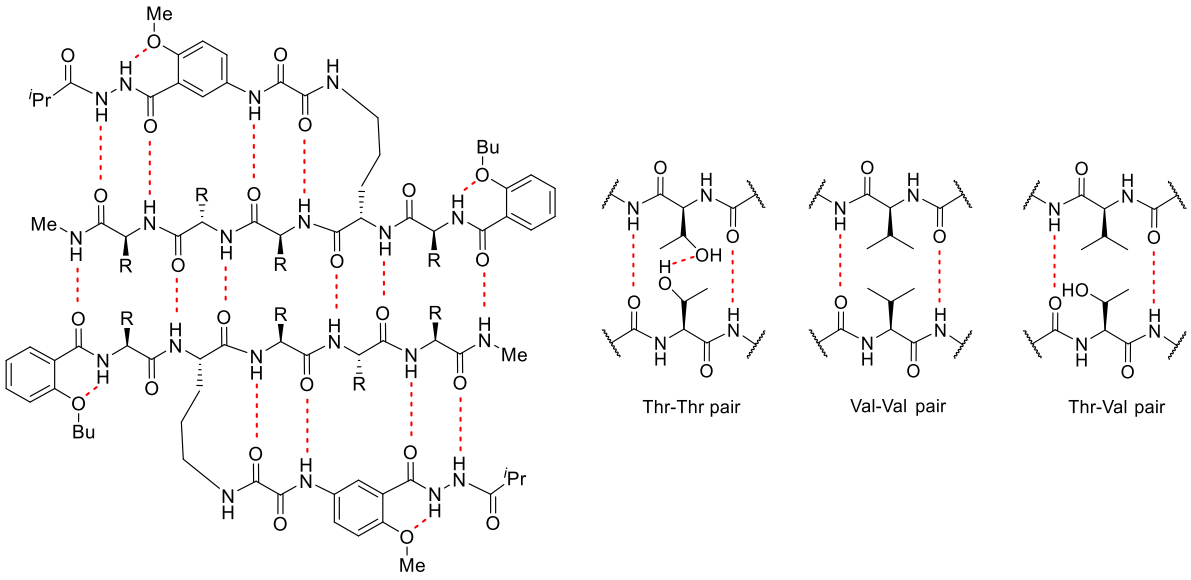


Figure 49. Homodimers synthesised by Nowick *et al.*,[176] including the amino acid pairs studied.

1.7.2 Therapeutic applications

Peptides containing unnatural β-turn mimics have been employed as therapeutics for a wide variety of applications, from the modulation of protein-protein interactions (PPIs)[177] to preventing amyloid aggregation,[178] among others. A few representative examples of their application within PPI inhibitors as well as within antimicrobial peptides are presented in this section.

1.7.2.1 Protein-protein interaction inhibitors

Protein-protein interactions are involved in most biological processes, and they are promising candidates in drug development. These interactions are challenging to inhibit with small molecules because their interfaces take place over a large surface and they involve a few hot-spot residues. In addition, many proteins are dynamic, undergo a conformational

change upon binding, or contain disordered regions. This has shifted the focus of much research towards designing PPI inhibitors based on their secondary structure.[93]

Many unnatural systems have been used to inhibit PPIs, namely cross-linked peptides and stapled systems, macrocycles containing hydrogen bond surrogates, among others.[136] Although most of the PPI inhibitors in the literature have α-helix conformations, a few β-hairpin-based PPI inhibitors have been reported which contain β-turn mimics. This approach can alter the conformation of the peptide structure, but depending on the mimic used this can be a positive effect as the desired conformation can be stabilised. Another potential benefit of using a β-turn mimic is the introduction of properties that improve the pharmacokinetic properties of the peptide.

The (D-Pro)-(L-Pro) template is one of the most widely used within cyclic β-hairpin peptides. It has been employed by Davidson *et al.*[179] to inhibit the interaction between the HIV-1 Tat protein and its transactivation response RNA, and by Fasan *et al.*[160] to inhibit the MDM2/p53 interaction. In the latter case the hairpin used acted as a mimetic of an α-helix, which held the relevant residues in the correct positions for interaction (Figure 50). This approach inspired the design of a family of cyclic β-hairpins that mimic the binding of the α-helical HIV-1 Rev protein to the Rev Response Element RNA.[180]

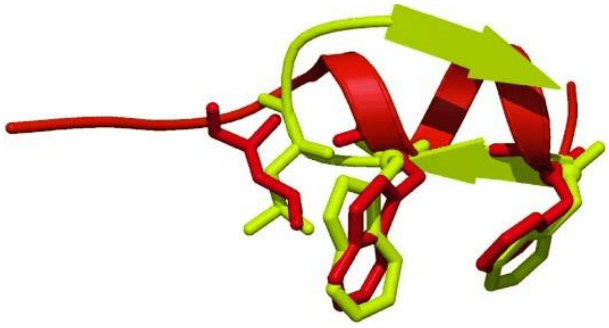


Figure 50. Superimposition of the p53 helical peptide (in red) and the β-hairpin mimic (yellow) designed by Fasan *et al*. Reproduced with permission from ref.[160]

Other β-turn mimics have been used in the literature – for example, Morse *et al.*[181] implemented ornithine within a cyclic β-hairpin to inhibit the interaction between the CDI toxin and an immunity protein (Figure 51). Rosenström *et al.*[182] introduced a bicyclic benzodiazepine-based β-turn mimic within the structure of angiotensin II and synthesised three different analogues that showed high binding affinity to the AT_2 receptor (Figure 51). In this case, the tyrosine side chain was introduced on the aromatic ring in the mimic, which seemed to increase binding.

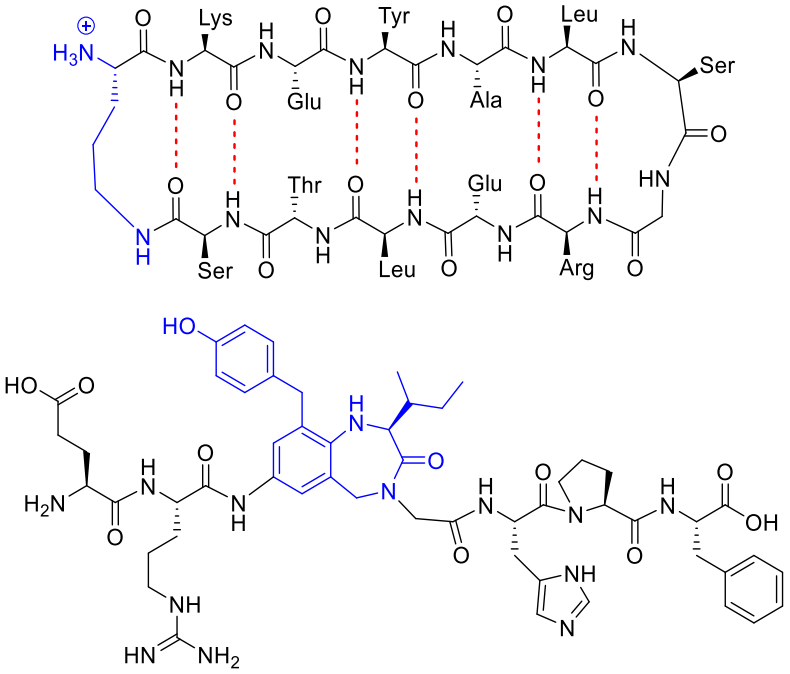


Figure 51. Top: ornithine-containing β-hairpin designed by Morse *et al.*[181] Bottom: peptide designed by Rosenström *et al.*[182] containing a benzodiazepine β-turn.

1.7.2.2 Antimicrobial peptides

The modification of β-turns has been exploited for the development of antimicrobial peptides.[162] Shankaramma *et al.*[159] used the (D-Pro)-(L-Pro) unit (Figure 52) and a xanthene-based template to synthesise new antimicrobial peptides based on the scaffolds of protegrin, polyphemusin and tachyplesin, three naturally occurring β-hairpin antimicrobial peptides. For this, they removed the disulfide bridges present in the native peptides and instead introduced one of the non-natural templates at one of the β-turns. The authors discovered a family of mimetics containing the (D-Pro)-(L-Pro) turn that showed potent broad spectrum antimicrobial activity.

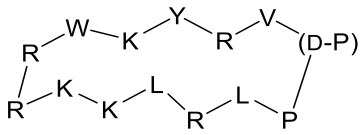


Figure 52. Representative structure of the family of mimetics synthesised by Shankaramma *et al.*[159]

The (D-Pro)-Gly turn unit has been used by Xu *et al.*,[133] who introduced it within a β-hairpin scaffold based on the tryptophan zipper motif. This gave rise to a family of five peptides with different lengths, out of which the 14-residue analogue presented the best antimicrobial activity.

1.8 Bioconjugation and ligation strategies

Bioconjugation refers to the site-specific covalent modification of biomolecules, which allows the incorporation of a new moiety that confers desirable properties to a protein. This strategy can take place between native amino acids and an unnatural moiety, as well as between two unnatural fragments. Bioconjugation is an area with a wide range of applications, from modifying peptide and protein therapeutics[183] to installing small molecule units that can act as probes to report on ligand binding, or as FRET reporters.[184]

There are three main requirements[185,186] that come into play when a bioconjugation method is being developed: firstly, the reaction should be biorthogonal. Secondly, the reaction conditions should be mild, ideally performed in aqueous media in order to preserve the integrity of the biomolecule. And thirdly, the bond generated should be stable under physiological conditions.

Many strategies[186] have been previously developed that use the side chains in native amino acids or the *C*- and *N*-termini as the bioconjugation site for a small molecule. These include the use of *N*-hydroxysuccinimide (NHS) esters on Lys residues (Figure 53) and the reductive amination between Lys and an aldehyde moiety, the use of maleimides (Figure 53) to react with Cys residues (to generate antibody-drug conjugates,[187] for example) as well as disulfide formation and exchange, among others.

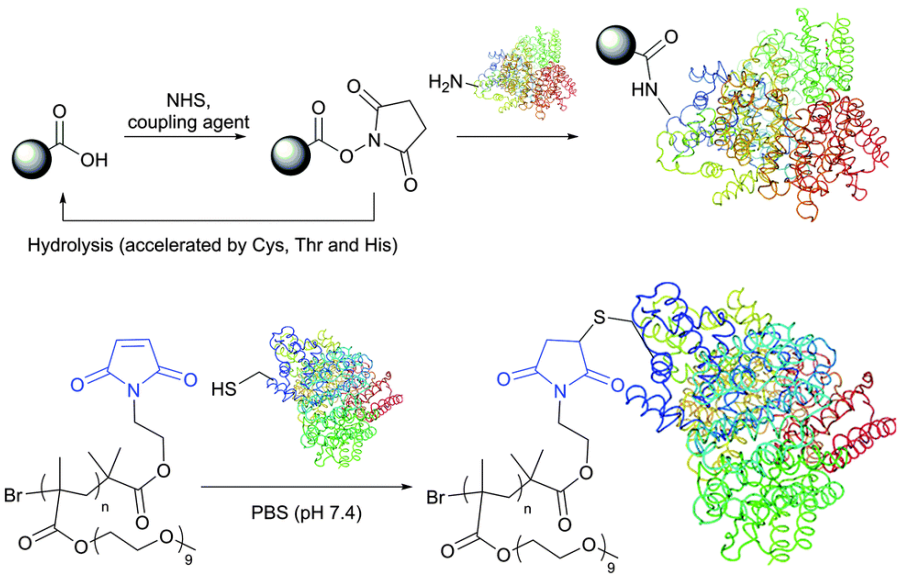
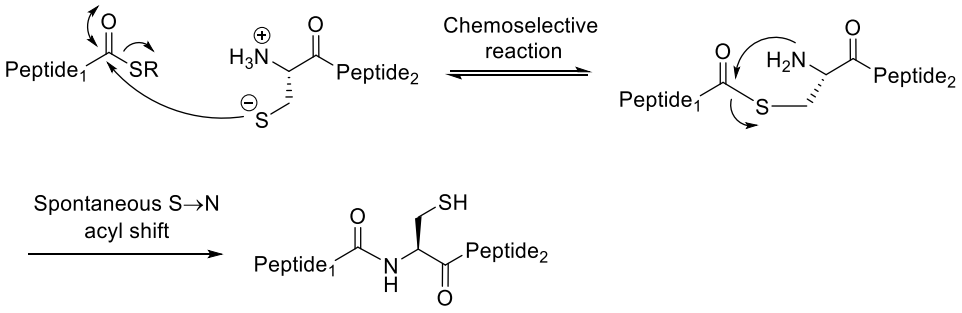


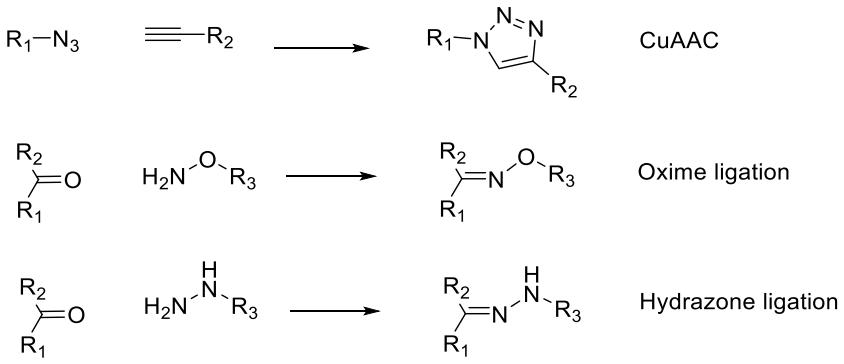
Figure 53. Top: bioconjugation between an NHS ester and a Lys residue in a protein. Bottom: bioconjugation between a maleimide-containing unnatural unit and a Cys residue in a protein. Reproduced with permission from ref.[186]

One of the key developments in the area was the native chemical ligation (NCL) by Dawson *et al*.,[188] where two peptides containing an *N*-terminal Cys and a *C*-terminal thioester can be ligated to generate a native peptide bond through an S→N acyl shift (Scheme 8). One of the main advantages of this strategy aside from its chemoselectivity is that the ligation site is a peptide bond, and no non-natural residual functionality is present in the final conjugate. Since its development, many variations on the synthesis have been introduced,[189,190] which include the acceleration of the ligation step by using thiol additives, the desulfurisation of Cys to generate an Ala residue after ligation and the use of selenocysteine. This is a versatile strategy that has found wide applications in the synthesis of peptides and proteins, peptide-oligonucleotide conjugates, as well as for the introduction of fluorescent probes onto proteins.[191]



Scheme 8. Mechanism of the native chemical ligation.

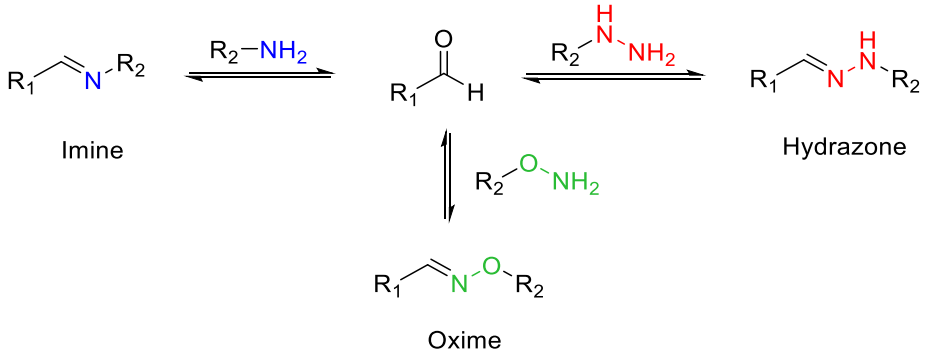
The use of canonical amino acids as the bioconjugation site has clear advantages, namely not having to modify the native peptide or protein prior to ligation. But there are limitations associated with this strategy, in particular the selectivity of the ligation step when more than one of the target amino acids are present. This has shifted the focus onto new ligation strategies (Scheme 9) where unnatural moieties are used on both ligation partners. Some of the most used strategies are click-type chemistries, like the copper(I)-catalysed azide–alkyne cycloaddition (CuAAC), and ligation chemistries that involve an aldehyde or ketone handle.[185]



Scheme 9. Copper(I)-catalysed azide–alkyne cycloaddition and oxime and hydrazone ligations.

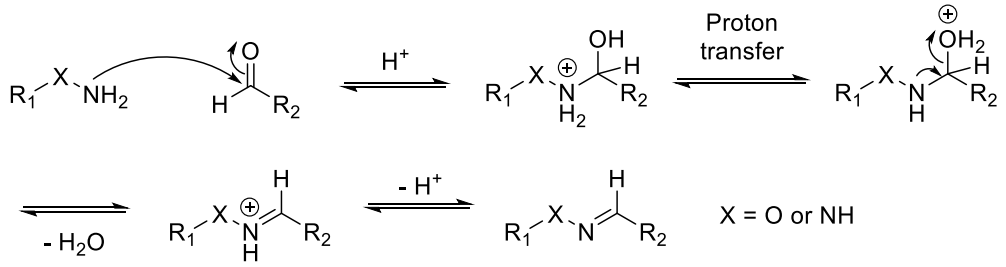
1.8.1 Hydrazone and oxime ligation

Hydrazone and oxime ligations (Scheme 10) are condensation reactions that take place between a ketone or an aldehyde and an α-nucleophile, namely a hydrazide, hydrazine or an *N*-alkoxyamine moiety, respectively. These are attractive reactions for bioconjugation as they take place in aqueous solution and involve a carbonyl functionality, which is not present in native proteins. Oxime ligation was first used for the synthesis of proteins by Rose in 1994.[192]



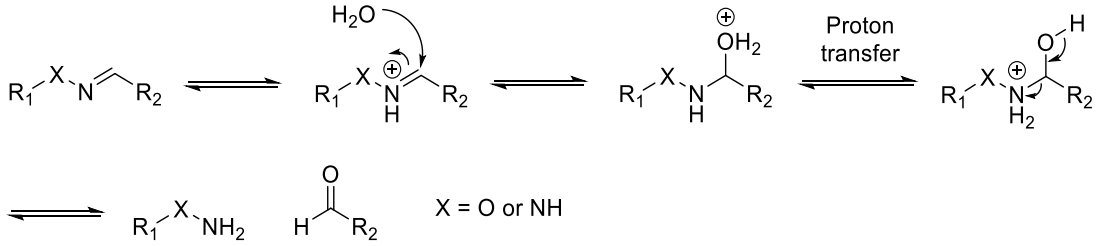
Scheme 10. Formation of imines, hydrazones and oximes from a carbonyl moiety.

These ligations are faster at above millimolar-range concentrations and at a pH of 4.5 (see mechanism in Scheme 11). The nature of the carbonyl moiety also affects the speed of the ligation: in general, the carbonyl group of an aldehyde is more electrophilic and reactive than that of a ketone, therefore ligations involving aldoximes tend to be faster and in consequence have been more explored in the literature than ketoximes.[193] Although the use of an acidic pH and an aldehyde can accelerate a ligation, in general these reactions are still slow. This has sparked interest in the development of nucleophilic catalysts to improve the reaction rates.[194,195]



Scheme 11. Mechanism of oxime and hydrazone ligation.

The stability of oximes and hydrazones is higher than that of imines due to the α-effect caused by the additional heteroatom, but these systems can undergo hydrolysis when they are in an acidic medium (Scheme 12). Kalia and Raines[196] studied the reaction rates of the hydrolysis of isostructural alkylhydrazones, acylhydrazones and an oxime at pH of 5 to 9 through ^{1}H NMR and found that oximes are much more stable than hydrazones due to the higher electronegativity of oxygen. This inductive effect reduces the basicity of the imine-like nitrogen, which in turn disfavours the collapse of the hemiaminal intermediate, which is the rate-determining step at the pH range under study.



Scheme 12. Mechanism of hydrolysis of oximes and hydrazones.

The stability of these systems can be enhanced by reducing the C=N bond with a reducing agent like sodium cyanoborohydride in order to generate the more stable hydrazide or *N*-alkoxy amine.

Oxime and hydrazone ligations have found multiple applications in the literature, from the cyclisation of peptides[197,198] to their use in dynamic combinatorial libraries.[193] Importantly, they have been used as mimics of structural elements: in the case of hydrazones, as peptide bond surrogates in dipeptides[199] and as hydrogen bond surrogates for an *i* to *i+4* hydrogen bond in an α-helix.[142] In the case of oximes, they have been used in the stabilisation of α-helices by means of oxime side-chain cross-links in collagen triple helices (Figure 54)[200] and a dimeric coiled-coil,[201] in both cases showing an increased stability of the system. They have also been employed as staples with *i* to *i+4* and *i* to *i+3* spacings in α-helices, which increased their stability compared to the control peptide.[202]

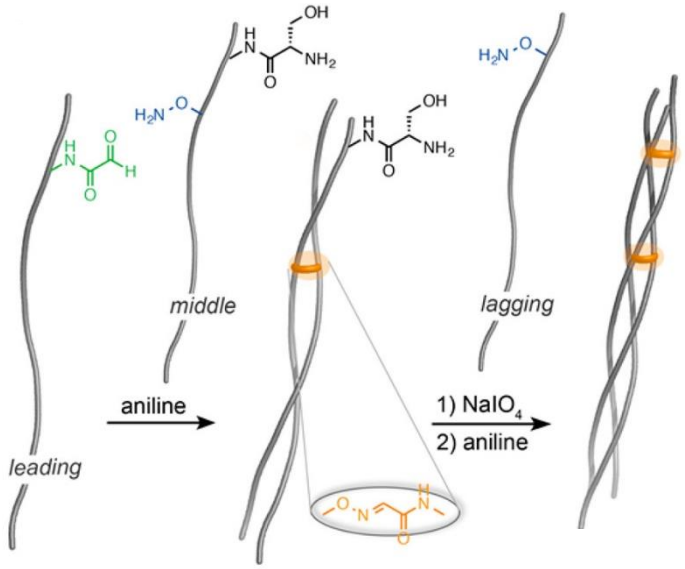


Figure 54. Oxime cross-linked triple helices designed by Hentzen *et al*. Reproduced with permission.[200]

1.9 Thesis aims

Proline isomerisation causes crucial structural changes within proteins that impact folding, conformation and ultimately function (see Section 1.4, topic further introduced in Chapter 2). The fact that proline and tertiary amide isomerisation more generally can initiate major conformational changes highlights how this scaffold has great potential to be used as the basis for a switch.

Although many switching systems have been previously designed, as introduced in Chapter 2, these are based on either small molecules or larger secondary structure changes. It was proposed that a switching system could be designed, which would exploit tertiary amide isomerisation in order to effect a conformational change within a short peptide backbone.

The design of β-turn mimics is an area that has been widely explored (introduced in Section 1.6.4), as they are a useful tool to control folding, provide higher stability and introduce new functionalities into a peptide. The most widely used β-turn mimics are based on D-amino acids, ornithine or triazole-containing systems, as they are easy to introduce in the peptide sequence, but these units lack structural preorganisation. The β-turns composed of completely unnatural units that have been previously reported do provide preorganisation, but they require the use of multi-step synthetic routes either to access the unnatural amino acids or to install the turn unit, which limits their applicability.

Furthermore, ligation chemistry is a promising strategy that has been successfully used in the bioconjugation and stabilisation of peptide systems (see Section 1.8.1). One of the drawbacks of ligation strategies is that, apart from the native chemical ligation reaction, the residual functionality at the ligation junction is left behind and serves no further purpose. Although oxime and hydrazone linkages have been proposed to act as peptide bond mimics, these units have not been used as mimics of elements of protein secondary structure yet. With this in mind, β-turn mimics were to be designed which would be accessed through one chemoselective ligation step, and that would incorporate the residual ligation functionality as part of the mimic unit.

The central aims of this thesis are:

- To understand the factors that influence conformational switching in small peptides containing proline and *N*-methylated alanine.

- To use hydrazide and oxime ligation as a means to install novel β-turn mimics within β-hairpin backbones and learn about their conformation and stability.

- To implement one of the designed β-turn mimics within a biologically relevant cyclic peptide. The possible changes in activity would be assessed and the applicability of the ligation strategy in cyclic systems would be evaluated.

2 Tertiary amide-based conformational switches

The work presented in this chapter has been performed in collaboration with Dr Drew Thomson and Miss Bethany Atkinson. Dr Drew Thomson generated the database and performed searches, and Miss Bethany Atkinson performed the molecular dynamics simulations.

2.1 Tertiary amide isomerisation and switching

2.1.1 The prolyl switch

Prolyl isomerisation has a considerable biological significance. It can be a rate-limiting step in protein folding,[116,203] which is slowed down by the rotation around the C-N bond and the formation of a twisted transition state. It has been found that isomerisation to the *cis* form in proteins alters a network of hydrogen bonds that is essential for folding to occur, which dramatically affects folding kinetics. It is also relevant in the regulation of protein function,[6,118,204,205] as its isomerisation results in various structural changes that can be local but also extend along the protein backbone;[206] and in signalling pathways,[207] for example in transmembrane signalling.[208]

The energetic difference between the *cis* and *trans* states for the prolyl bond is smaller than that for secondary amide bonds. In *cis* secondary amide bonds, the C^{α} atoms are very close to each other (about 3 Å, as found in a database survey by Jabs *et al.*[209]), which imposes a higher steric strain than in the *trans* configuration. This causes *cis* secondary amide bonds to only be found in 0.028% of the Xaa-nonPro bonds (where Xaa is any amino acid) in the PDB. Conversely, in prolyl bonds the two states have a similar steric strain and therefore the *cis* configuration is more commonly observed (5.2% of the Xaa-Pro bonds in the PDB).[92]

2.1.1.1 Local sequence effects

The occurrence of a *cis* prolyl bond is highly influenced by the local sequence.[5] This has been explored by several database screenings and experimental projects, which aimed to elucidate what sequence-based rules determine the conformation of a prolyl bond.

An experimental study by Reimer *et al.*[210] focused on peptides of the sequence Ac-Ala-Xaa-Pro-Ala-Lys-NH_2 through NMR and found that the amino acids that caused a higher fraction

of *cis* Pro were Trp (37.7%), Tyr (42.0%) and Phe (23.0%). Notably, a correlation was observed between the presence of a *cis* prolyl bond and the occurrence of a type VI β-turn.[211]

Pal *et al.*[5] screened a set of non-redundant proteins from the PDB and found that 5.7% of Xaa-Pro bonds were *cis*, of which more than 10% had a Trp in the Xaa position. The authors proposed the formation of a CH-π interaction with the H^{α} or H^{δ} of Pro as an explanation for the high occurrence of aromatic residues preceding it. Indeed, Tóth *et al.*[68] found that 68.3% of Xaa–(*cis*)Pro fragments in the PDB contained CH–π interactions, compared to only 16.8% of the Xaa–(*trans*)Pro fragments.

Other screenings of the PDB[71] have confirmed that aromatic-Pro sequences are the most likely to adopt a *cis* bond, and analyses of the geometries between both residues have shown that the Pro H^{α}, H^{β} and H^{δ} are commonly under the aromatic ring, confirming the presence of a CH-π interaction. Chakrabarti *et al.*[212] continued their studies on this interaction and found that unlike the CH-π interaction between two aromatic rings, where the edge to face geometry is required, Pro can stack onto an aromatic ring while getting involved in the CH-π interaction through the H^{α} or H^{δ} position (Figure 55), in a face to face geometry. This way, the aromatic-Pro interaction is stabilised by both the CH-π interaction and the hydrophobic effect.

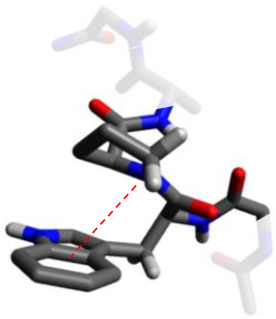


Figure 55. C-H···π interaction between Pro H^{δ} and Trp.

Although the most influential position is Pro-1 (the residue preceding Pro), positions Pro-2 and Pro+1 (the residue following Pro) can also affect isomerisation. These were examined by Zondlo *et al.*[213] for xyPz peptides and they found that although most residues in the Pro-2 position had a moderate effect on isomerisation, aromatic residues relatively disfavoured the *cis* configuration. In the Pro+1 position, aromatic residues as well as Asn, Ala, Val and Pro promoted the *cis* configuration. This is commensurate with previous observations that the second Pro in Pro-Pro-aromatic sequences also has a high propensity to adopt a *cis* prolyl bond.[214,215]

2.1.2 *N*-methylated residues

N-methylated residues are another interesting moiety that could be used in the construction of a switch. Like proline, *N*-methylated residues cause the energy difference between the *cis* and *trans* isomers to be small, so *cis* peptide bonds are also frequently observed. Peptides with this modification are very important in the regulation of biological functions, from targeting the cell membrane to binding DNA and RNA.[216] Recent research has put the spotlight on *N*-methylated amino acids in the area of medicinal chemistry as a means to increase the metabolic stability,[217] affinity and selectivity of bioactive peptides.[218,219]

Although *N*-methylation has been widely explored in drug development, far fewer studies have been performed around its effects on peptide conformation. The studies to date in this area have focused on introducing *N*-methylation within cyclic peptides and in combination with D-amino acids in order to favour the formation of a β-turn. For example, Laufer *et al.* found that, in combination with D-Ala, (*N*-Me)-Ala could serve as a substitute for Pro within cyclic poly-Ala pentapeptides.[220]

The effects of *N*-methylation on backbone conformation have two components: the removal of its hydrogen-bonding capability and the steric hindrance generated.[216] This causes its impact on conformation to be long-range, not only affecting the modified residue but neighbouring amino acids as well. In larger peptides the hydrogen bond donor component becomes more important, whereas the steric component is more prominent in smaller systems. Steric hindrance takes place about the *N*-methylated peptide bond but also between the methyl group and the amino acid's side chain, which affects *cis*/*trans* configuration. The influence of the *N*-methylated residue's side chain in conformation was studied by Laufer *et al.*[221] in the peptides c(RGD(D-F)K), c(RGD(D-F)(*N*-Me)-K) and c(RGD(D-F)(*N*-Me)-V). They found that the non-methylated parent peptide contained all *trans* amide bonds, while (*N*-Me)-K introduced 15% of *cis* configuration and (*N*-Me)-V introduced no *cis* configuration at all. This proved the important role of the side chain in defining backbone conformation, and showed how β-branched amino acids introduce more steric hindrance in the *cis* isomer.

Lahiri *et al.*[107] used *N*-methylation as a strategy to design stable β-hairpins through modification of the turn sequence. They synthesised two libraries of peptides, with β-turns composed of a D-residue in position *i+1* and an L-*N*-methylated residue at position *i+2*. The first library was a D-residue scan with an (*N*-Me)-Ala in position *i+2* and the second was an

(*N*-Me)-residue scan using D-Ala in position *i+1*. They found that the β-hairpin resulting from the combination of the best residues from each library (a (D-V)(*N*-Me)-R turn) was exceptionally stable. This confirmed that the two positions in the turn act in synergy to modulate folding and that both residue side chains and the steric hindrance from *N*-methylation are important factors in defining backbone conformation.

2.1.3 Disulfide bridges: two-state behaviour

Two-state behaviour has been observed in peptides containing disulfide bridges. For example an antimicrobial peptide called Bactenecin forms a type I β-turn in the disulfide-bridged form but when it is in the reduced form its conformation varies depending on the environment lipophilicity.[222] In addition, disulfide-bridged peptides have a tendency to form β-turns and many contain proline within the ring.[223,224] This has led to an interest in the study of proline isomerisation within disulfide-bridged peptides.

Rabenstein *et al*.[225] studied the reduced and oxidised forms of peptides of the sequences Ac-Cys-Gly-Pro-Cys-NH_2 and Ac-Cys-Pro-Xaa-Cys-NH_2 (where Xaa = Phe, His, Tyr, Gly and Thr, Figure 56) and characterised the kinetics of Pro isomerisation. They found that the *cis-trans* isomerisation rates were faster in the disulfide-bridged sequences than in the free thiol ones by a factor ranging from 4 to 13. This suggested that the presence of Pro in the disulfide ring brought an unusual conformational flexibility to the system. Molecular mechanics calculations indicated the mechanism for this rate acceleration, whereby the disulfide constrains the backbone and facilitates the formation of a hydrogen bond between the NH proton of the Pro-Xaa bond to the proline nitrogen lone pair in the transition state. This lowers the energy barrier to access the transition state from the *cis* configuration by 5.86 kJ/mol.

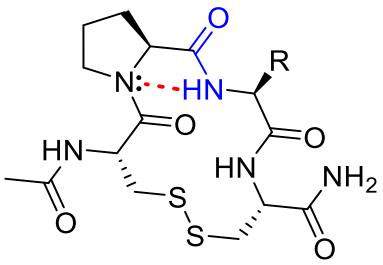


Figure 56. Transition state of the *cis-trans* isomerisation of Pro in the Ac-Cys-Pro-Xaa-Cys-NH_2 peptide studied by Rabenstein *et al*.[225], where the Pro-Xaa peptide bond is represented in blue and the hydrogen bond in red.

Further work by Rabenstein *et al.*[226] focused on the *cis* configuration content in peptides of the sequence Ac-Cys-Pro-$(Xaa)_n$-Cys-NH_2 (where n = 0-4), both in the dithiol and disulfide forms (Table 3). It was observed that at pH 3 the *cis* configuration population was 10-20% for the dithiol sequences, but for the disulfide peptides the population varied over a wider range, namely 9% to 77%.

Table 3. Peptides studied by Rabenstein *et al.*,[226] including their *cis* Pro content.

Peptide sequence	% *cis* Pro
Ac-Cys-Pro-Cys-NH_2	10.3
Ac-Cys-Pro-Cys-NH_2 disulfide bridged	21.7
Ac-Cys-Pro-Phe-Cys-NH_2	12.0
Ac-Cys-Pro-Phe-Cys-NH_2 disulfide bridged	9.3
Ac-Cys-Pro-Phe-Ala-Cys-NH_2	14.9
Ac-Cys-Pro-Phe-Ala-Cys-NH_2 disulfide bridged	59.5
Ac-Cys-Pro-Phe-Ala-Ala-Cys-NH_2	13.3
Ac-Cys-Pro-Phe-Ala-Ala-Cys-NH_2 disulfide bridged	12.2
Ac-Cys-Pro-Phe-Ala-Ala-Ala-Cys-NH_2	20.4
Ac-Cys-Pro-Phe-Ala-Ala-Ala-Cys-NH_2 disulfide bridged	76.9

Interestingly, a dependence was observed between the *cis* content and the size of the disulfide-bridged ring formed. The *cis* content increased compared to the free thiol peptide in disulfide rings with an odd number of amino acids, whereas it decreased when the ring contained an even number of residues. This is the first report showing that the conformation distribution in Pro-containing disulfide-bridged peptides can be controlled through loop size, and raises the question of whether these systems could be used to promote switch-like behaviour.

2.2 Designed peptide-based conformational switches

Molecular switches are key in many biological processes, from the control of translation and transcription to signal transduction[227]. Recently, synthetic molecular switches have been at the forefront of chemical biology due to their great potential to be used as sensors and

regulatory elements. Many switchable proteins have been designed[227] to obtain control of various cellular signalling processes and also to engineer complex synthetic signalling networks. Smaller systems have also been developed, which encompass both small molecule-based switches as well as peptide-based secondary structure changes. Larger scale secondary structure changes are the most widely studied types of peptide switches,[228] as they include α-helix to β-sheet transitions and folded to unfolded state changes.

One of the key elements in the design of a switch is its method of actuation. In this section, peptide-based conformational switches in the literature will be reviewed according to their control element.

2.2.1 Ligand binding

The use of a ligand to induce conformational switching upon binding is a common way of actuation, which can include binding to a receptor as well as a metal. For example, Mueller and Grossmann[229] have described a conformational switching system composed of two coiled-coil helices. When the peptide is unbound it adopts a closed conformation, but upon binding to a receptor through the helix-connecting region it forms an open conformation (Figure 57).

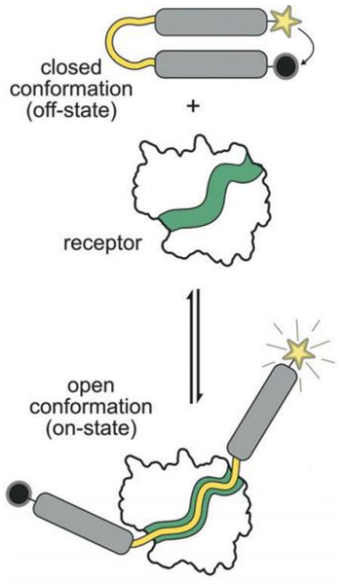


Figure 57. Hairpin to open conformation switch by Grossman *et al*. Reproduced with permission from ref.[229]

2.2.2 Phosphorylation

Phosphorylation is a post-translational modification that plays an important part in signal transduction events. Phosphorylated Ser/Thr-Pro motifs act as binding surfaces for many proteins,[230] including Pin1, a prolyl isomerase.[231]

Phosphorylation was used as a stimulus for a conformational switch in a system designed by Signarvic *et al.*,[232] where an unstructured monomer switches into an antiparallel four-helix bundle upon phosphorylation (Figure 58). The introduction of the post-translational modification stabilises the helical conformation by establishing new interactions within the tertiary structure, which causes the energy gap to be bigger between the two states, therefore making the switch viable.

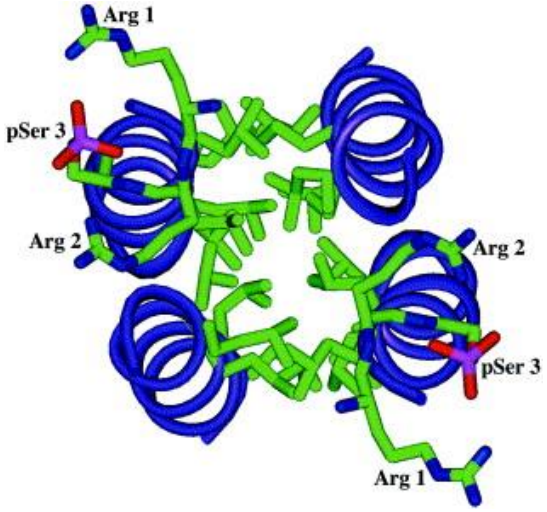


Figure 58. Four-helix bundle studied by Signarvic *et al*. Reproduced with permission from ref.[232]

2.2.3 Photo-switches

Photo-controlled peptide switches have especially been explored within both linear[233,234] and cyclic[235] β-hairpin backbones, with the use of azobenzene as the switching moiety[134,236] being the most widespread. This linker is usually introduced in place of the original turn sequence in order to control the formation of an ordered structure. One particular example is the photo-controlled β-hairpin designed by Dong *et al.*,[135] which is based on the introduction of an azobenzene-based moiety in place of the turn sequence in a TrpZip β-hairpin designed by Cochran *et al.*[80] (Figure 59). When the chromophore is in the *trans* state the peptide has a disordered, irregular structure. Upon isomerisation into the *cis*

configuration the azobenzene unit induces folding of the peptide into a β-hairpin structure, therefore acting as a β-turn mimic.

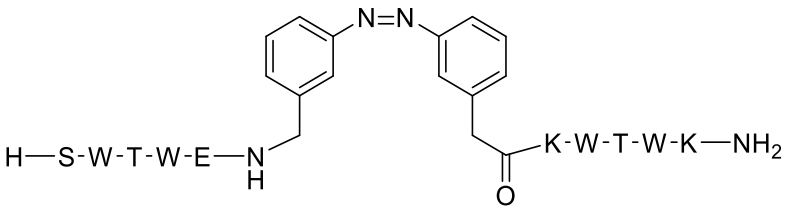


Figure 59. Photo-controlled β-hairpin designed by Dong *et al*.[135] in the *cis* configuration.

2.2.4 pH changes

There are various examples in the literature of designed conformational switches actuated by a change in pH. For example, Anderson and Andersen[237] designed the β-turn sequence HPATGK, which acts as a pH switch within a β-hairpin backbone. When at pH 2.5, the sequence is unstructured but at pH 8 it becomes highly structured and stable. The use of this turn sequence within a WW domain protein showed that depending on where the sequence is introduced in the protein, it will affect the system's stability to a different extent. When placed in the nucleating β-hairpin, the effect of the turn sequence is global, whereas if placed in another region, the effect will be limited to a local region (Figure 60). This effect is caused by the protonation of a His residue within the turn sequence, therefore most of the loss in stability takes place at pH from 3 to 6.5.

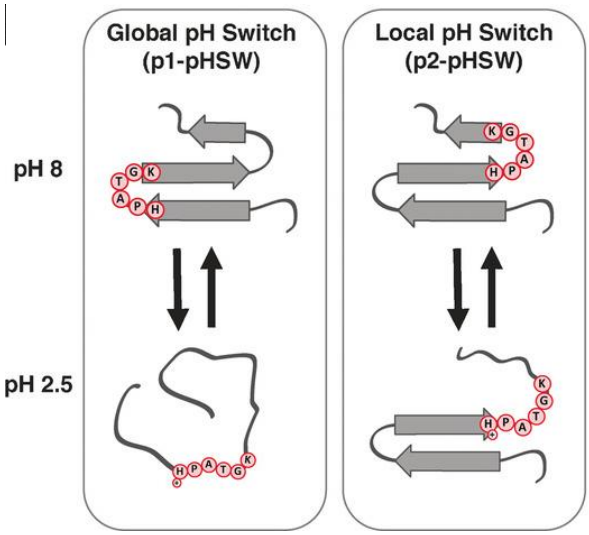


Figure 60. Left: global pH switch induced by the turn sequence being placed at the nucleating β-hairpin. Right: local pH switch. Reproduced with permission from ref.[237]

2.2.5 Solvent changes

Conformational switching induced by solvent changes has been explored, where most systems rely on the hydrogen-bonding capability of the solvent. Awasthi *et al.*[238] designed an 8-residue peptide with an Aib-Gly β-turn motif that adopts a β-hairpin conformation when in hydrogen-bonding solvents like DMSO and methanol due to the solvation of the backbone C=O and NH groups. When dissolved in non-hydrogen bonding solvents, the peptide adopts a helical conformation. Similar results are observed in the work by Balamurugan *et al.*,[239] where a heterochiral peptide containing non-natural amino acids adopts a helical conformation in chloroform and an extended conformation in DMSO.

2.2.6 Redox systems

One of the earlier studies regarding redox switches was performed by Gellman *et al.*[240] They explored the oxidation of methionine residues as the stimulus for a conformational change in an 18-residue peptide, which adopts an α-helical conformation when reduced and a β-strand conformation upon oxidation.

Disulfide formation has been explored by Woolfson *et al.*,[241] who designed a system with a monomeric helical hairpin structure in the disulfide-bridged form. Upon reduction of the disulfide, the system switches to a coiled-coil dimer, although it is a one-way switch and it cannot return to its starting conformation.

A more subtle conformational change is induced in the cyclic peptidomimetic designed by Jiang *et al.*, (Figure 61)[242] which contains a thioether that adopts a type I β-turn structure upon oxidation to the sulfone form, due to the formation of a hydrogen bond between NH(*i*+*2*) and O=S.

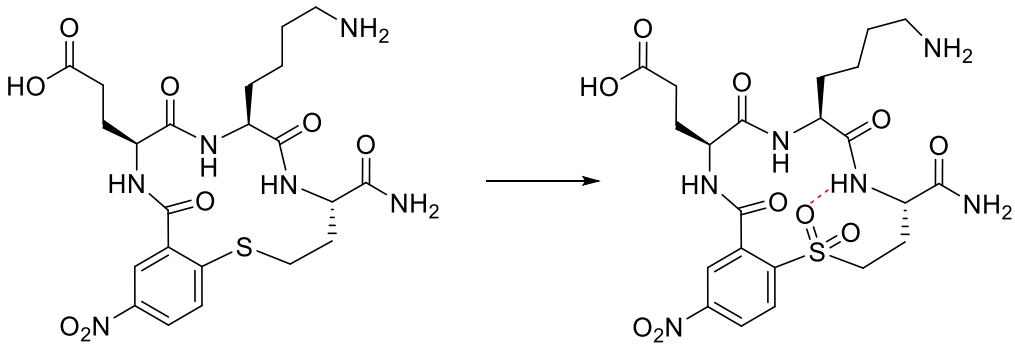


Figure 61. Analogues studied by Jiang *et al.*[242] The hydrogen bond between NH(*i*+*2*) and O=S is shown in red.

The formation of a disulfide bridge has been shown to promote β-turn formation in CxxC motifs[243] and a switching system has been observed in a native antimicrobial peptide,[222] where a type I β-turn is formed in the oxidised form but no turn is observed in the reduced form. Similar observations apply for vicinal disulfides: Carugo *et al*.[38] performed a database search and found that the presence of a vicinal disulfide correlated with the formation of a tight turn in the protein. They also observed that very different conformations were adopted in the reduced and oxidised states. Vicinal disulfides have been used to modulate the enzymatic activity of Ribonuclease A by mutating cysteine residues into the native sequence, so that when the disulfide bridge is formed the system's activity is decreased and it can be recovered upon reduction.[244]

2.3 Hypothesis and chapter aims

There are many molecular switches in nature that control biological processes. These systems have inspired the design of a wide variety of switchable proteins to further understand the biological process of interest and how switching takes place,[232] as well as to act as biosensors.[245] Most of these systems are based on either small-molecule scaffolds or larger, protein secondary structure changes, with very few of them utilising a smaller peptide scaffold.

The prolyl switch is central to the control of folding, function and signalling in proteins[118] and proline isomerisation has been proven to cause significant conformational changes (see Sections 1.4 and 2.1). Hence there was an opportunity to design a switching system based on a short peptide backbone that would make use of proline isomerisation and undergo a conformational change.

With this in mind, there was potential for the design of smaller switching systems where tertiary amide isomerisation could be exploited. A system was envisioned where tertiary amide isomerisation would act as the conformational difference between two states. A backbone of the type Xaa-Pro-Yaa was devised, where the effects of residues Xaa and Yaa on proline conformation were to be evaluated. *N*-Methylated alanine was to be explored as well, as an alternative to proline. This would provide information that would aid the design of a switching system actuated through disulfide formation.

To accomplish this design, the following aims were defined:

- Study of the local sequence dependence of proline conformation within the backbone of interest, in particular the cooperativity between the x and y residues.

- Replacement of proline by *N*-methylated alanine, in order to evaluate the difference in *cis* and *trans* populations.

- To design disulfide-actuated switches.

- Assessment of the switches through NMR spectroscopy.

2.4 Xaa-Pro-Yaa double mutant cycle

When designing a conformational switch it is key to choose a backbone where both conformations are energetically accessible. Proline-containing sequences are known to adopt *cis* and *trans* configurations (this nomenclature will be used in the context of tertiary amides hereinafter), but the occurrence of the more sterically hindered *cis* configuration is highly dependent on local sequence, causing its population to vary considerably. In this section, a database was searched to identify potential switching loops with the motif Xaa-Pro-Yaa and their *cis* conformer content. The chosen unconstrained sequence was to be studied through a double mutant cycle (DMC), using NMR to identify the conformations adopted and their populations.

2.4.1 Database screening and sequence selection

Computational methods were used in this project as a fast and pragmatic way of exploring the protein sequence space in order to select target sequences for experimental evaluation. A database was generated by Dr Drew Thomson which contained high-resolution, non redundant protein crystal structures. This database was then filtered to extract all the loop regions in proteins that were flanked by regions of secondary structure. This was based on the idea that, unlike other regions of proteins where the local conformation can be affected by non-local interactions, the conformation of a loop should be dominated by local interactions. This loop subset was used to identify Xaa-Pro-Yaa motifs and their *cis* conformer content, in order to study the influence of local sequence on the occurrence of *cis*

Pro. This search was performed considering a *cis* Pro if the ω angle for the prolyl bond was 0 ± 20° and *trans* if ω was 180 ± 20° (Figure 6, Figure 62).

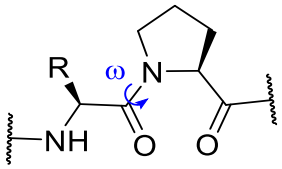


Figure 62. Omega dihedral angle in Pro.

A list of the loops containing this motif was generated, which included the sequence and *cis* content for the Pro residue. A heat map showing the *cis* Pro content for all possible combinations of Xaa and Yaa in Xaa-Pro-Yaa is presented in Figure 63. It was observed that the sequences with the highest *cis* content featured an aromatic residue in the position preceding Pro, had an aromatic amino acid in both Xaa and Yaa positions, or contained a Pro-Pro motif.

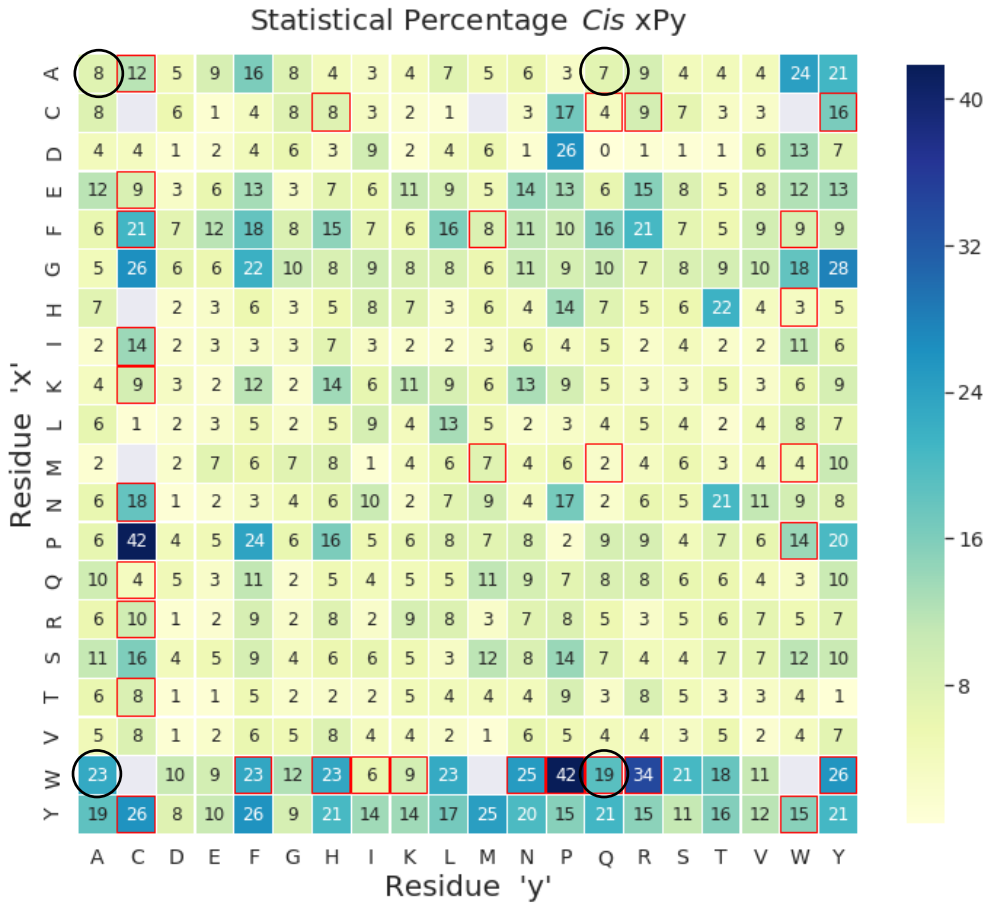


Figure 63. Heat map showing the results from the database search for Xaa-Pro-Yaa sequences, with the *cis* Pro content values (%) for each combination of amino acids. The residue combinations that had a total number of hits lower than 50 are marked with red boxes, and the combinations with less than 20 hits are not included (coloured grey). The results circled in black correspond to the sequences chosen for analysis.

Inspired by these results, the sequence Ac-GWPQG-NH_2 **1** was chosen for experimental evaluation. The reasons for this were twofold: firstly, aromatic residues preceding Pro have been shown to stabilise the *cis* configuration, and secondly, a polar residue was preferred in the Yaa position in order to improve the solubility of the peptide. The peptide was to be flanked by Gly residues and capped in both termini to avoid end-effects, as the free charged *N*- and *C*-termini could establish additional interactions and therefore affect the conformation adopted by the peptide.[246] The sequence chosen did not have the highest *cis* Pro content observed in the database search – this was in order to be able to observe a change in the proportion of *cis* Pro present upon mutation. There is a balance between finding a sequence for which the *cis* configuration is accessible, and one where the *cis* content is so high that a mutation will not cause an observable change.

This sequence was to be studied through a double mutant cycle (DMC), where two mutations would be introduced in order to calculate the free energy of interaction between Trp and Gln (Figure 64). Alanine was chosen as the non-mutated residue and therefore, the four sequences Ac-GWPQG-NH_2 **1**, Ac-GAPAG-NH_2 **2**, Ac-GWPAG-NH_2 **3** and Ac-GAPQG-NH_2 **4** were selected for experimental evaluation.

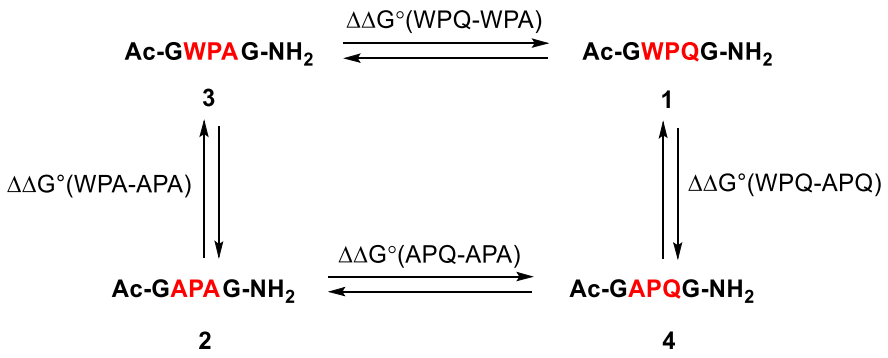


Figure 64. Double mutant cycle for Ac-GWPQG-NH_2, using Ac-GAPAG-NH_2 as the non-mutant sequence.

The individual ΔG° between *cis* and *trans* states for every peptide could be calculated using Equation 11, which in turn would be used to obtain the ΔΔG° for the double mutant cycle (Equation 12).[79]

$$\Delta G^\circ = -RT\ln K$$

Equation 11. Calculation of the free energy for each sequence, where R is the gas constant (8.31 $JK^{-1}mol^{-1}$), T is the temperature (298 K) and K is the equilibrium constant.

$$\Delta\Delta G° = \Delta G(00) - \Delta G(0j) - \Delta G(i0) + \Delta G(ij)$$

$$= \Delta G(APA) - \Delta G(APQ) - \Delta G(WPA) + \Delta G(WPQ)$$

Equation 12. Calculation of the free energy for the DMC.

The absolute error arising from the ΔΔG° calculation could be calculated using Equation 13, which would account for inaccuracies in the integration of the NMR signals:

$$\varepsilon^a(a+b+c+d) = \sqrt{\varepsilon^a(a)^2 + \varepsilon^a(b)^2 + \varepsilon^a(c)^2 + \varepsilon^a(d)^2}$$

Equation 13. Calculation of the absolute error (ε^a) for ΔΔG°, where a to d correspond to the individual ΔG° for each peptide.

Where:

$$\varepsilon^a(G) = \frac{RT}{K}\varepsilon^a(K)$$

Equation 14. Calculation of the individual absolute errors of the ΔG° for each peptide.

2.4.2 Molecular dynamics simulations

Metadynamics simulations were run on the chosen DMC by Miss Bethany Atkinson, in order to predict the behaviour of the selected peptides regarding proline isomerisation. Metadynamics is a powerful enhanced sampling technique that attempts to improve the timescale problem of MD. In this technique, the system is described by collective variables (CVs), which are parameters chosen to describe the system's coordinates in space. In order to describe Pro isomerisation, ψ and the improper dihedral angle ζ were chosen as the CVs for these simulations (Figure 65).

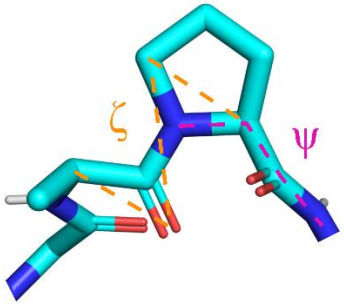


Figure 65. Proline ψ and ζ angles.

The simulations were run for the four peptides (at 298 K), which afforded the *cis* Pro content for each sequence as well as the energy difference between the *cis* and *trans* states (Table 4).

Table 4. *Cis/trans* ratio, *cis* content and free energy difference between the *cis* and *trans* states for the four peptides studied.

Sequence	*Cis/trans* ratio	*Cis* content (%)	ΔG° (kJ/mol)
Ac-GWPQG-NH_2, **1**	0.46	31.6	1.87
Ac-GAPAG-NH_2, **2**	0.09	8.2	5.85
Ac-GWPAG-NH_2, **3**	0.39	27.9	2.36
Ac-GAPQG-NH_2, **4**	0.1	8.7	6.13

These models showed that the sequences with a tryptophan in the position preceding proline had a substantially higher occurrence of the *cis* configuration, which was in accordance with the literature.[5,210] The ΔΔG° for the double mutant cycle was calculated as shown in Equation 15, and the value of -0.07 kJ/mol was obtained. The negative value of the ΔΔG° indicated that there was positive cooperativity taking place between tryptophan and glutamine. This finding was evaluated against experimental studies.

$$\Delta\Delta G^\circ = \Delta G(\mathrm{APA}) - \Delta G(\mathrm{APQ}) - \Delta G(\mathrm{WPA}) + \Delta G(\mathrm{WPQ})$$

Equation 15. Calculation of the free energy for the DMC.

2.4.3 Peptide synthesis and NMR studies

The peptides were synthesised *via* standard microwave-assisted SPPS, using the Fmoc/*tert*-butyl (^{t}Bu) strategy and HCTU activation on TentaGel S RAM resin (0.24 mmol/g) in order to obtain amidated *C*-termini. The *N*-terminus of the peptides was acetyl capped using acetic anhydride to avoid end-effects.

The sequences were studied through NMR, which allows the identification of both conformations as prolyl isomerisation is slow compared to the NMR time scale. Intending to choose conditions that mimicked physiological conditions, experiments were initially run in PBS buffer, but the peptides were found to precipitate gradually from solution due to the

buffer's high salinity. Instead, potassium phosphate (20 mM, at pH 7.4) was used for the samples, which were at concentrations of 100 or 200 μM in 5% D_2O (at 298 K).

Two-dimensional NOESY experiments were used to identify both conformers: the diagnostic NOE cross peaks for the *cis* configuration take place between H^{α}(Pro-1) and H^{α}(Pro), whilst the ones for the *trans* configuration are between H^{α}(Pro-1) and H^{δ}(Pro) (Figure 66). The two-dimensional NMRs for **4** were analysed, and it was confirmed that the major conformation corresponded to the *trans*. As all the sequences in the double mutant cycle were very similar, it was assumed that the *trans* configuration was the major one for the remaining three peptides. This extrapolation was backed by the literature, as there have been previous studies of sequences of the type Ac-GxPG-NH_2 that have shown that the *trans* configuration is the most populated.[211]

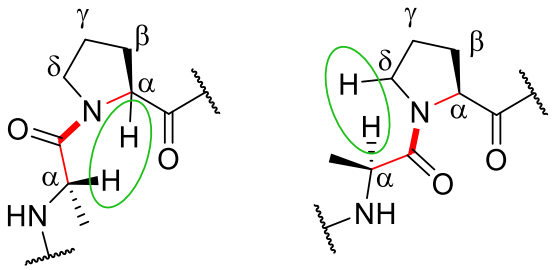


Figure 66. Diagnostic NOE signals for the *cis* (left) and *trans* (right) conformations.

The population of each conformation was determined through integration of the peaks that had clear signals for both conformers in the ^1H NMR experiments. This provided a *cis/trans* ratio, which can be considered equivalent to the equilibrium constant between states in equilibrium. The values found are presented in Table 5.

Table 5. *Cis/trans* NMR ratios (at 298 K) and ΔG° for the four sequences in the DMC, as well as ΔΔG° for the system.

Sequence	NMR *cis/trans* ratio	*Cis* content (%)	ΔG° (kJ/mol)
Ac-GWPQG-NH_2, **1**	0.35	26%	2.60 ± 0.07
Ac-GAPAG-NH_2, **2**	0.09	8.3%	5.71 ± 0.12
Ac-GWPAG-NH_2, **3**	0.43	30%	2.09 ± 0.06
Ac-GAPQG-NH_2, **4**	0.08	7.4%	6.26 ± 0.01
ΔΔG° (kJ/mol)			**- 0.04** ± 0.16

The sign of the ΔΔG° obtained indicated that the cooperativity between Trp and Gln was slightly positive, which meant that both residues stabilised the *cis* configuration to a greater degree than they did individually. The significance of this cooperativity effect might not be large, especially when comparing the small value of the ΔΔG° to the uncertainty of the measurement.

Another interesting observation was that there was a difference of an order of magnitude between the *cis* content of Trp-containing sequences and non-aromatic ones. The higher population of the *cis* configuration could be explained by the presence of a stabilising interaction between the Pro pyrrolidine ring and the Trp indole, namely a CH-π interaction involving the Pro H^{α} or H^{δ}.[68,212] This interaction was observed in the MD models, which took place when **3** and **1** adopted the *cis* Pro conformation (Figure 67). When in the *trans* configuration, a more extended conformation was observed, where Trp was in closer proximity to the backbone instead of the pyrrolidine ring.

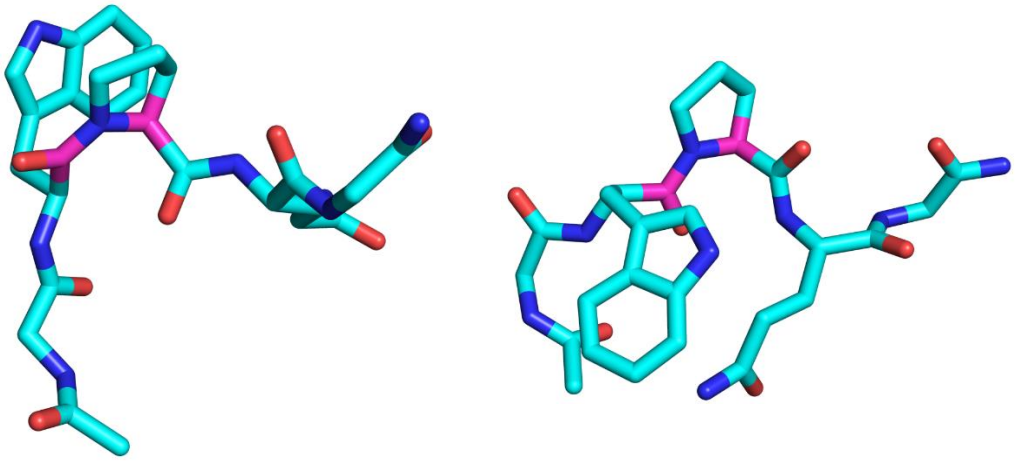


Figure 67. MD models for 1 with *cis* Pro (left) and *trans* Pro (right). The prolyl bond dihedral is represented in pink.

The values of ΔG° and *cis* content for each peptide obtained from the MD simulations were comparable to the ones from the NMR studies, which validated metadynamics as a method to choose sequences for experimental evaluation and predict their behaviour. The biggest difference between values was for **1**: this could be caused by a poor representation of the side chain interactions between tryptophan and proline in the simulations. The AMBER forcefield used in these simulations represents atoms as point charges, and therefore does not account for the resonance and the quadrupole moment of π systems.[247] This could have a significant effect on the conformations predicted, as the accuracy of the force field becomes very important in smaller systems.[248]

This double mutant cycle study confirmed that the presence of an aromatic residue preceding proline favours the occurrence of the *cis* configuration in unconstrained peptides and highlighted the cooperativity between residues in the local sequence. This established a foundation on which a switching system could be based.

2.5 *N*-methylated analogue 5

N-methylated residues are an interesting alternative to proline, as they are also tertiary amides but they provide more flexibility to the system, which might favour a higher population of the *cis* configuration. The peptide Ac-GW(*N*-Me)AQG-NH_2 (**5**, Figure 68) was studied in order to compare its *cis* configuration content with its proline equivalent, **1**.

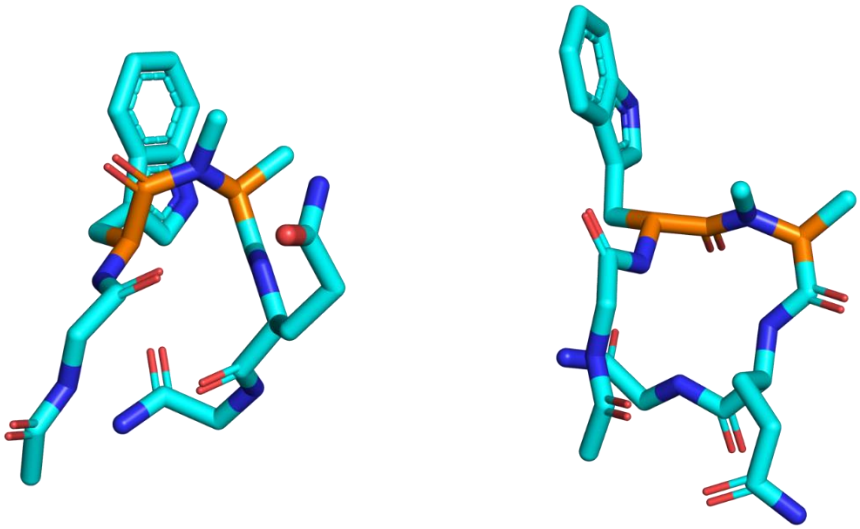
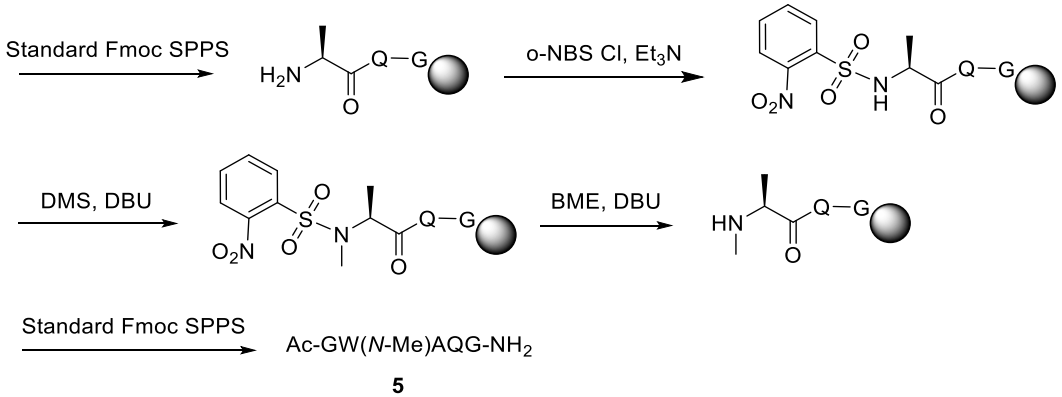


Figure 68. Energy-minimised *cis* (left) and *trans* (right) conformers of 5. The atoms involved in the ω dihedral angle are shown in orange.

5 was synthesised through standard microwave-assisted SPPS, using the Fmoc/^{t}Bu strategy and *N*,*N*′-diisopropylcarbodiimide (DIC)/Oxyma activation on an Fmoc-Rink Amide AM resin (0.74 mmol/g). A double coupling was performed for the tryptophan following the *N*-methylated residue, as the higher steric hindrance on the secondary amine causes the coupling to be more difficult.

The *N*-methylation step was performed using *o*-nitrobenzenesulfonyl chloride (*o*-NBS) activation followed by methylation using dimethyl sulfate and DBU (Scheme 13). This protocol was adapted from the strategy initially described by Miller and Scanlan,[249] which was later optimised by Chatterjee *et al*.[219,250] In this work, instead of collidine, triethylamine was used as the base in the *o*-NBS protection step, and DMF was used as a solvent instead of NMP during the *o*-NBS deprotection.



5

Scheme 13. Synthetic route towards 5.

This sequence was studied by NMR (1 mM, in potassium phosphate buffer 200 mM, both in 5% D_2O and 100% D_2O at 298 K) to find that the *cis/trans* ratio was 0.66 (40% *cis* content). The identity of the major and minor conformations was confirmed through NOE cross peaks (Figure 69), where an NOE between H^{α}((*N*-Me)-Ala-1) and the *N*-Me group indicated a *trans* configuration, and one between H^{α}((*N*-Me)-Ala-1) and H^{α}((*N*-Me)-Ala) indicated a *cis* configuration. It was found that the major conformation corresponded to the *trans* form of the peptide. The experiments in 100% D_2O were used to visualise the H^{α} and the *trans* configuration NOE. In addition, a cross peak between the *N*-methyl groups in both conformations with the opposite sign of the NOE cross peaks was observed, confirming that exchange between conformations was taking place.

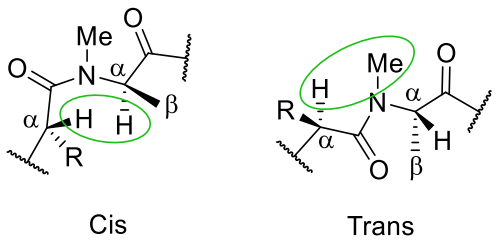


Figure 69. Expected diagnostic NOE cross peaks for the *cis* and *trans* configurations in 5.

Comparing to its proline equivalent, **1**, the *cis* content for the *N*-methylated peptide was higher (40% for **5** and 26% for **1**). This was consistent with the literature: Marraud *et al.*[251] found that *N*-methylated Ala favoured the *cis* configuration more than Pro in dipeptides. The authors suggest that this could be due to the higher flexibility of the *N*-methylated residue, which is less restricted along the ϕ dihedral angle and therefore allows the system to access more geometries compared to its Pro counterpart. This could favour the formation of a CH-π interaction between (*N*-Me)-Ala and Trp, which could stabilise the *cis* configuration. An alternative explanation could be the polarisation of the *N*-methyl group in (*N*-Me)-Ala: if the partial charge on the methyl group was more positive than the equivalent H^{δ} position in proline, that would establish a stronger CH-π interaction.

The NOESY experiments were examined to find more information: for the major conformation, cross peaks were observed between the Ala *N*-methyl group and Trp $H^{\varepsilon 3}$/Trp $H^{\zeta 3}$ and Trp $H^{\delta 1}$. For the minor conformation they were observed between the Ala *N*-methyl group and Trp $H^{\delta 1}$. This could indicate the presence of a stabilising CH-π interaction in both conformations.

Energy minimisation experiments were run using the GAFF forcefield and a conjugate gradients method for **5** and **1**, both in the *cis* and *trans* configurations. The models showed that (*N*-Me)-Ala or Pro were in close proximity to Trp only when the peptides adopted the *cis* configuration. Namely, the H^{α} for (*N*-Me)-Ala in **5** was at 2.3 Å from the indole ring (Figure 70) and the H^{α} for Pro in **1** was at 2.2 Å from it, suggesting that a CH-π interaction could take place.

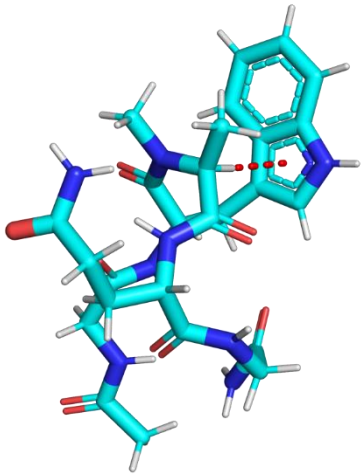


Figure 70. Energy-minimised model of 5 with *cis N*-Me-Ala, showing the possible CH-π interaction between (*N*-Me)-Ala H^{α} and Trp.

2.6 Disulfide-actuated switches

Intramolecular disulfide formation was chosen as a potential actuation mechanism for the design of an initial switching system based on Pro isomerisation. The conformation of Pro has been shown to be affected by disulfide formation,[252] and CxyC tetrapeptides have been shown to promote β-turn formation in the oxidised form (where x and y were Ala or Gly).[243] A system was envisioned with a sequence of the type CxPC (Figure 71), where an open conformation would exist in the resting state, with reduced cysteines and *trans* Pro. Upon oxidation, enough energy would be provided for the isomerisation to *cis* Pro to take place, which would give rise to a closed peptide conformation. This was inspired by the type VI β-turn geometry, where a *cis* Pro is present at the *i+2* position. Because the cooperativity between Trp and Gln in the series of peptides studied in Section 2.4 was not very significant, it was hypothesised that Gln could be substituted by one of the Cys residues. Gly residues were added to frame the peptide, in order to avoid end-effects and resemble a longer protein environment. Proline was to be tested initially, and (*N*-Me)-Ala would be explored if the designs were successful.

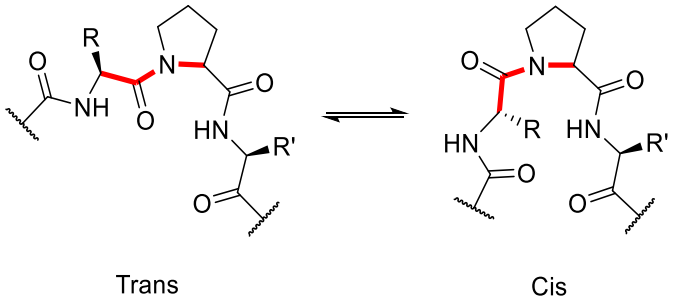


Figure 71. *Trans* (left) and *cis* (right) Pro rotamers corresponding to the open and closed conformations in a CxPC system, respectively.

The sequences Ac-GCWPCG-NH_2 (**6**) and Ac-GCAPCG-NH_2 (**7**) were chosen for study by analogy with the double mutant cycle from Section 2.4. Energy minimisation models were generated for both *cis* and *trans* configurations of the reduced peptides, which showed that an open conformation was adopted by both peptides with a *trans* Pro. With a *cis* Pro a more compact conformation was observed where the Cys thiol groups were in close proximity, which could facilitate the formation of the disulfide bridge (Figure 72).

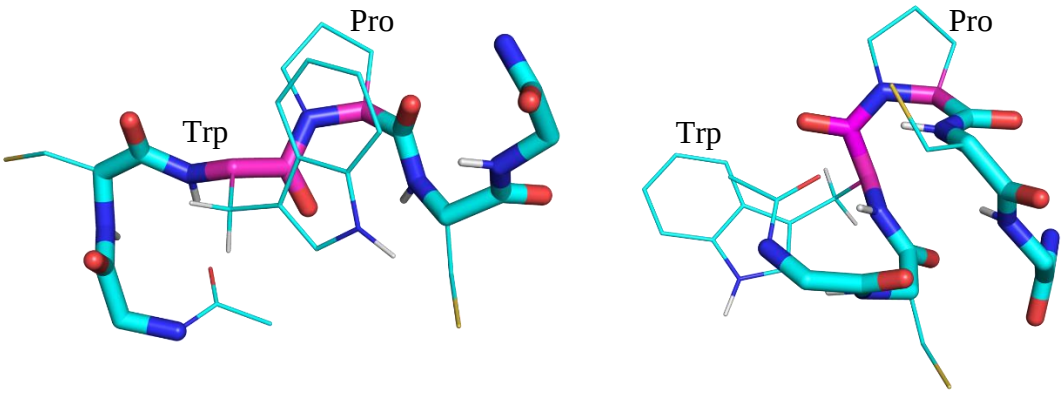


Figure 72. *Trans* (left) and *cis* (right) conformations of 6. The ω dihedral angle is represented in pink.

Both sequences, in the reduced and oxidised forms (Figure 73), would be synthesised and analysed through NMR to ascertain how *cis* Pro content changes as a function of the Cys oxidation states, whilst comparing the effects of Trp and Ala in equivalent systems.

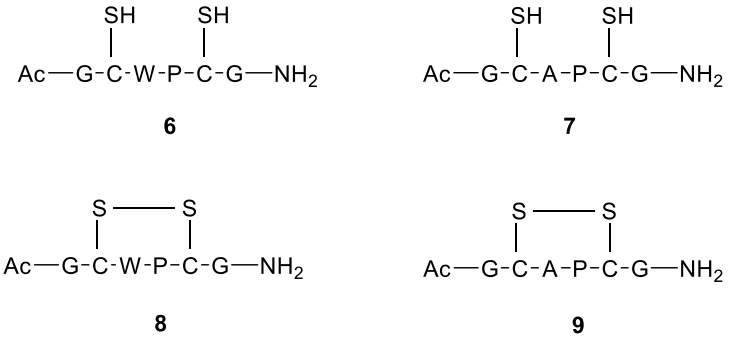


Figure 73. The four sequences chosen for evaluation.

2.6.1 Database search

A database search was performed by Dr Drew Thomson to explore in what contexts disulfide bridges are present and how these could be used as an actuation method for a CxPC-type switch. To this end, the loop database introduced in Section 2.4.1 was filtered to obtain sequences of the type CxPC, the omega (ω) angle of Pro as well as information on the oxidation state of the Cys residues. Only ten sequences were found, therefore not much information could be extracted regarding the relationship between Pro conformation and the oxidation state of the cysteines. Thus, the chosen systems would be evaluated experimentally to ascertain if the formation of the desired *i* to *i+3* disulfide bridge is coupled to a geometry change.

2.6.2 Synthesis

6 and **7** were synthesised through standard microwave-assisted SPPS, in order to obtain good coupling efficiency,[253] using the Fmoc/^{t}Bu strategy and HCTU activation on a TentaGel S RAM resin (0.24 mmol/g) to afford the *C*-terminal amide. In order to prevent epimerisation, cysteine was coupled using a method at 50°C. The peptides were purified through RP-HPLC and the disulfide formation was then performed in solution.

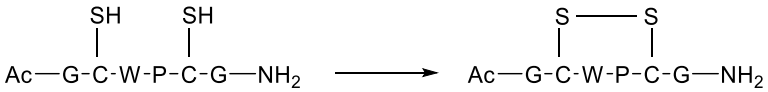
A method that didn't result in oligomerisation was sought for the disulfide formation reaction. An initial oxidation test performed on **6** using conditions by Calce *et al*.[254] (2.5 mM peptide in NH_4HCO_3 0.1 M, with air bubbled through the sample) caused a white solid to precipitate during the reaction. When the sample was analysed by LCMS no product or oligomer mass could be observed, suggesting that the high concentration caused intermolecular disulfide formation and precipitation. A second air oxidation test was performed using a peptide concentration of 0.1 mM (entry 1, Table 6), which was close to

completion after 43 h (92% yield estimated by LCMS) and afforded the desired disulfide-linked peptide as the major product, although minor dimer formation was observed (5% by LCMS).

In order to decrease the reaction time, three more tests were performed on **6** at 0.1 mM concentration. Following conditions by Albericio *et al.*,[255,256] a test was performed using *N*-chlorosuccinimide (entry 2, Table 6), which can accelerate disulfide formation when used in slight excess in aqueous solution. The starting material was consumed in 15 minutes but the reaction showed 10% conversion to the dimer (estimated yield by LCMS) and it was very time sensitive, as if left for a longer time the product disappeared and instead, other unidentified by-products were generated.

A further reaction test was performed using 15% volume of DMSO in NH_4HCO_3 0.1 M (entry 3, Table 6), which reached completion in 6 h by LCMS but generated 10% of the dimer. Dipyridyl disulfide was explored as well (entry 4, Table 6), employing 3 equivalents of DPDS in MeOH/ NH_4HCO_3 0.1 M (conditions by Maruyama *et al.*[257]). The reaction was complete within an hour by LCMS analysis and only generated traces of dimer.

Table 6. Oxidation methods screened for 6. All reactions were performed at room temperature. The yields were estimated *via* LCMS. For entry 1, 3% of starting material was left unreacted. Decomposition was observed for entry 2 if the reaction was left for longer than 0.25 h. For entry 3, 2% of starting material was left unreacted. Complete conversion was observed for entry 4 after 1 h.



6 **8**

Entry	Oxidation method	Solvent	Reaction time	Estimated yield of product	Estimated yield of dimer
1	Air oxidation	NH_4HCO_3 0.1 M	43 h	-	5%
2	NCS	MeCN/water	0.25 h	90%	10%
3	DMSO	NH_4HCO_3 0.1 M	6 h	90%	10%
4	DPDS	NH_4HCO_3 0.1 M	1 h	95%	Traces

It was determined that the DPDS-mediated oxidation was the most efficient, as it proceeded in short reaction times and did not cause significant dimerisation. After performing the oxidation using the chosen conditions and purifying the peptides *via* RP-HPLC, **8** was obtained in a 56% yield and **9** in a 50% yield.

2.6.3 NMR analysis

For the reduced-Cys sequences, two-dimensional NMR experiments (COSY, TOCSY, NOESY, ROESY) were run at 1 mM concentration in potassium phosphate buffer (200 mM at pH 7.4) and 5% D_2O, at 298 K. To the samples was added 1 mM DTT-d_{10} in order to keep the cysteines in the reduced form. The *cis* and *trans* configurations were identified as minor and major, respectively, through the NOE cross peaks between the Pro H^{δ} and Trp or Ala H^{α} (for *trans*) and the cross peak between the Pro H^{α} and Trp or Ala H^{α} (for *cis*).

^{1}H NMR experiments were run for the disulfide-bridged peptides in the same conditions, without DTT and at 100 (for **9**) and 200 μM (for **8**) concentrations due to limited material. The NMR experiments for the reduced peptides were also run at these lower concentrations and showed no difference from the 1 mM concentration. This indicated no concentration-dependent change in conformation or aggregation.

Two sets of signals corresponding to the *cis* and *trans* Pro configurations were observed for the disulfide-bridged peptides. The chemical shifts of the bridged peptides were comparable to the non-bridged ones, thus the major conformation was identified as the *trans* Pro by analogy. This was necessary because the peptides tumbled in between regimes, which caused the appearance of NOE cross peaks to be limited. This prevented the complete assignment of the peptide backbone. The closely related sequence Ac-CAPC-NH_2 has been studied in the literature,[252] for which the *trans* was the major conformation observed as well. No additional sets of signals were observed arising from the disulfide bridge diastereomers, which could be due to them fast exchanging or due to the presence of only one diastereomer.

The *cis* Pro content was calculated for the four peptides from the integrals obtained in the ^{1}H NMR experiments (Table 7).

Table 7. *Cis/trans* ratio and *cis* content for the four peptides.

Sequence	*Cis/trans* ratio	Percentage of *cis* configuration
Ac-GCWPCG-NH_2 thiol, **6**	0.43	30%
Ac-GCWPCG-NH_2 disulfide, **8**	0.49	33%
Ac-GCAPCG-NH_2 thiol, **7**	0.26	21%
Ac-GCAPCG-NH_2 disulfide, **9**	0.34	25%

Although a 3-4% increase in the occurrence of *cis* Pro was observed, this was not a significant magnitude, therefore indicating that the formation of the disulfide bridge did not force the system into the *cis* configuration. This showed that the system was not working as the intended conformational switch.

Consistent with the literature and what was observed in Section 2.4, the Trp-containing peptides had ~10% more *cis* Pro, highlighting the role of this residue in stabilising that conformation.

Similar sequences have been studied by Sui and Rabenstein[252] to explore the occurrence of *cis* and *trans* Pro in a wide range of disulfide-bridged peptides and their corresponding dithiol equivalents. Among others, the tetrapeptide Ac-CAPC-NH_2 was investigated, for which a *cis* content of 46% was found as the disulfide-bridged sequence and an 8% *cis* as the free thiol form. These values differ from the ones found for **7** and **9**, which could be due to the Gly flanking the peptides but could also be affected by the different conditions used to run NMR experiments (a pH of 3 was used by Sui *et al*. whereas pH 7 was used in this work).

2.7 Vicinal disulfide systems

2.7.1 Hypothesis and database searches

A new switch design was envisioned where the formation of a vicinal disulfide would confer a constraint that would induce conformational switching. Significant conformational changes caused by vicinal disulfides have been previously observed in native proteins, as introduced in Section 1.2.3.1. A Yaa-Xaa-Pro-Cys-Cys system was devised, where the Xaa

position would be a *cis* Pro-inducing amino acid, like Trp. In the reduced state a *cis* Pro would take place, which would induce the formation of a closed conformation (similar to a type VI β-turn). Upon formation of a vicinal disulfide bridge, an open conformation would be adopted with a *trans* Pro as the major isomer (Figure 74). The formation of the disulfide bridge would be incompatible with the presence of a *cis* Pro due to the increased strain in the backbone.

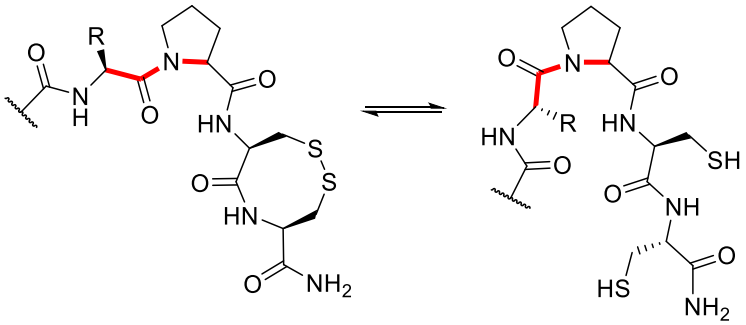


Figure 74. Two states of the Yaa-Xaa-Pro-Cys-Cys switch envisioned, with a *trans* (left) and *cis* Pro (right).

It was envisioned that only *cis* Pro would accommodate the formation of a closed conformation, while the constraint introduced by the vicinal disulfide would cause the *trans* Pro state to adopt an open conformation. Energy-minimised models of the sequence Ac-TWPCC-NH_2 showed that the Cys side chains in the *cis* Pro state pointed in opposite directions (Figure 75), which would make the formation of the vicinal disulfide bridge more difficult than in the *trans* Pro state.

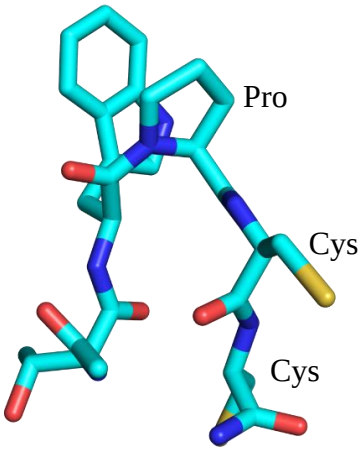


Figure 75. Energy-minimised model of Ac-TWPCC-NH_2 with reduced Cys and *cis* Pro.

This pronounced difference between states was also observed by Carugo *et al.*,[38] who performed a search for all sequences containing vicinal disulfides in the PDB and found that

very different conformations were adopted in the oxidised and reduced states. In the first case, a turn conformation took place and the Cys side chains pointed towards the same side. In the latter, an extended backbone was observed while the Cys side chains were on opposite sides of the main chain.

Sequences of the type xPCC were searched for in the loop database introduced in Section 2.4.1 by Dr Drew Thomson, but not much information could be obtained due to the low number of hits. Therefore, a rational design approach was adopted to choose the peptides for study. The first sequence chosen was Ac-TWPCC-NH_2 (**10**). Tryptophan was introduced preceding Pro because of the analogy to the other sequences studied in this chapter, as well as its proven ability to stabilise the *cis* Pro conformation. Thr was introduced as a polar amino acid, and because it could potentially aid the NMR assignment due to the characteristic high chemical shift of the H^{β} and the presence of the H^{γ} signals. The peptide Ac-TW(*N*-Me)ACC-NH_2 (**11**) was also chosen as a target for study due to the higher *cis* content observed for (*N*-Me)-Ala, as demonstrated in Section 2.5.

2.7.2 Synthesis and NMR studies

10 and **11** were synthesised through standard microwave-assisted SPPS, using the Fmoc/^{t}Bu strategy and DIC/Oxyma activation on an Fmoc-Rink Amide AM resin (0.74 mmol/g). Cysteine was coupled using a method at 50°C to prevent epimerisation. The peptides were purified and the disulfide formation was then performed in solution.

Disulfide formation was performed with the conditions used for the CxPC sequences (Section 2.6) but at 50°C, as vicinal disulfides are more constrained systems. This afforded **12** and **13** in a 43% and a 22% yield after purification, respectively.

Two-dimensional NMR experiments (COSY, TOCSY, NOESY) were run for each peptide at 1 mM concentration with 1 mM DTT-d_{10} for **10** and **11**, in potassium phosphate buffer (200 mM, pH 7.4) and 5% D_2O, at 298 K. For both peptides the major conformation was identified as the *trans* through the NOE cross peaks between Pro H^{δ} and Trp H^{α} for **10** and between Trp H^{α} and Ala *N*-Me for **11** (Figure 76).

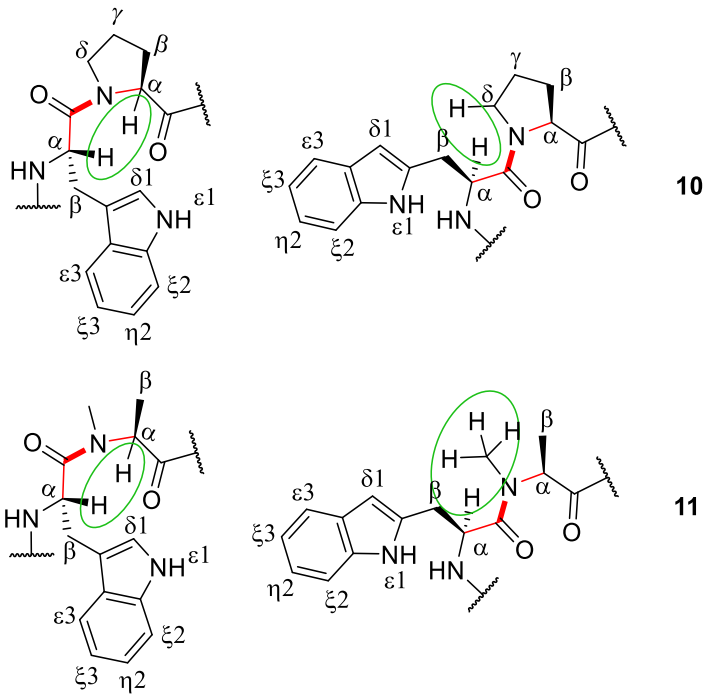


Figure 76. Diagnostic *cis* (left) and *trans* (right) NOE cross peaks (in green) for 10 and 11.

The spectra for the disulfide-bridged **12** and **13** were more complex than initially expected, as more than the expected two conformations were observed. In the case of **13**, five different NH signals were present in the Trp $H^{\varepsilon 1}$ region at 298 K (Figure 77). This could be due to the disulfide rotamers, which could be exchanging slowly and therefore become visible in the NMR timescale. Another possibility could be the occurrence of a *cis* amide bond in the *C*-terminal cysteine (as has been observed previously in the literature[258]).

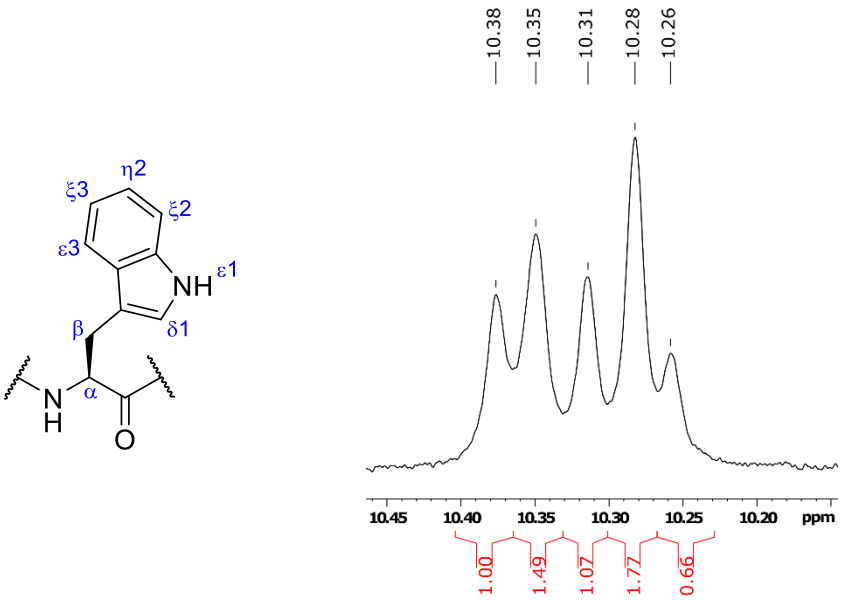


Figure 77. ^{1}H NMR of 13 showing the multiple conformations for Trp H$^{\varepsilon 1}$ at 298 K, and labelling of the Trp protons. ^{1}H NMR experiment performed at a 1 mM peptide concentration, in potassium phosphate buffer (200 mM, pH 7.4) and 5% D_2O, at 298 K on a 600 MHz spectrometer.

Temperature-dependent ^{1}H NMR experiments were run for **13** from 298 K to 278 K every 5 degrees to ascertain whether the conformations were interconverting and facilitate assignment. As the temperature decreased the backbone amide NH exchange with solvent was slowed down[259] and new signals appeared, although the Trp H$^{\varepsilon 1}$ signals stayed the same. Therefore it was unlikely that the conformations observed were in exchange within the NMR timescale (Figure 78).

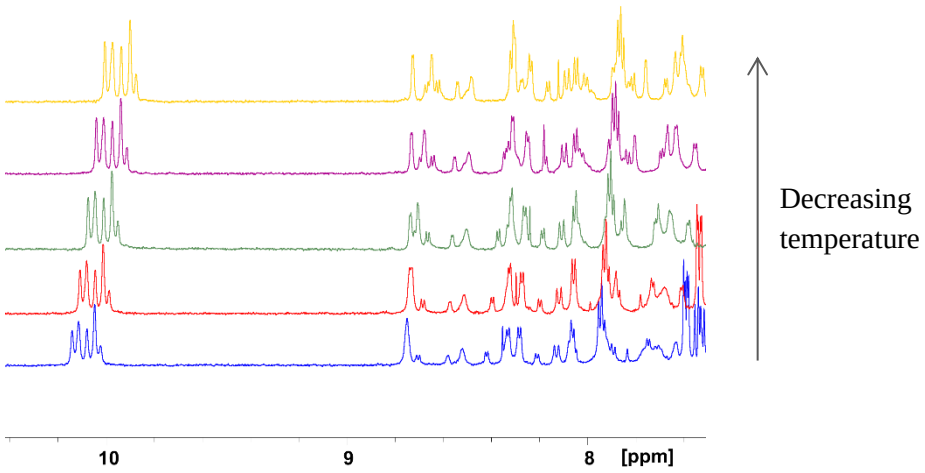


Figure 78. From bottom to top: 298, 293, 288, 283 and 278 K experiments for 13. ^{1}H NMR experiments performed at a 1 mM peptide concentration, in potassium phosphate buffer (200 mM, pH 7.4) and 5% D_2O on a 600 MHz spectrometer.

In order to facilitate assignment, experiments were run in DMSO-d_6. It was hoped to obtain a solubility improvement and clearer spectra, as water suppression would not be an issue. 1D ^{1}H NMR experiments were run at 298, 303, 308 and 313 K, heating up the sample instead of cooling it down due to the viscosity of DMSO (Figure 79). The experiment at 303 K showed the best peak resolution for the Trp $H^{\varepsilon 1}$ protons, although a slightly different pattern was observed compared to the experiments in buffer: in this case, five different conformations were still observed, but a different proportion of the signals and an overlap between two of the peaks took place.

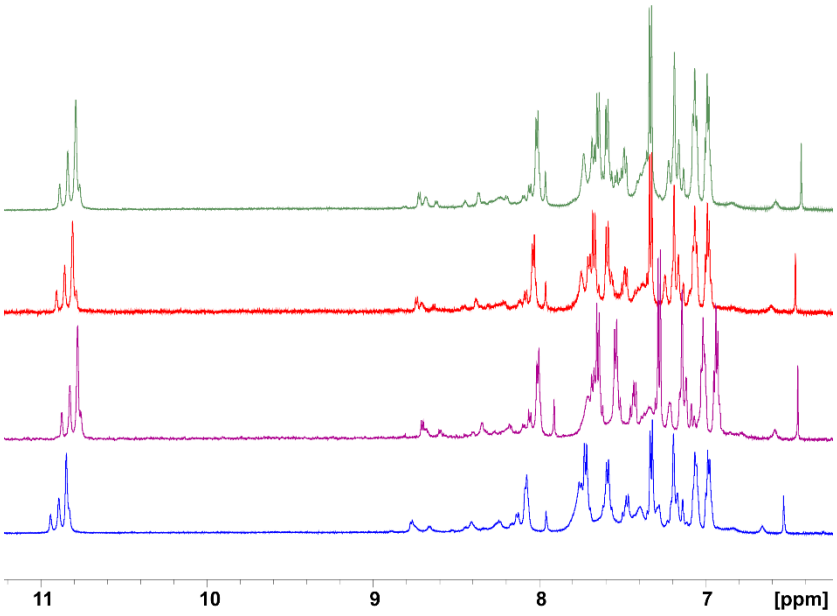


Figure 79. From bottom to top: 298, 303, 308 and 313 K experiments for 13 in DMSO-d_6 (1 mM peptide concentration, 600 MHz spectrometer).

A ROESY experiment was used for assignment instead of a NOESY at 303 K, as the latter showed only exchange cross peaks of the opposite sign to the diagonal. The ROESY experiment confirmed the exchange peaks observed and showed more correlations in the aliphatic to backbone NH region. Assignment for this system was very complex, as all conformations overlapped considerably in chemical shift. The α, β, δ1 and ε1 signals for the tryptophan residue in all five conformations were assigned, which allowed the identification of the *cis* or *trans* configuration of the system and associate that with the five $H^{\varepsilon 1}$ signals (Figure 80) for integration.

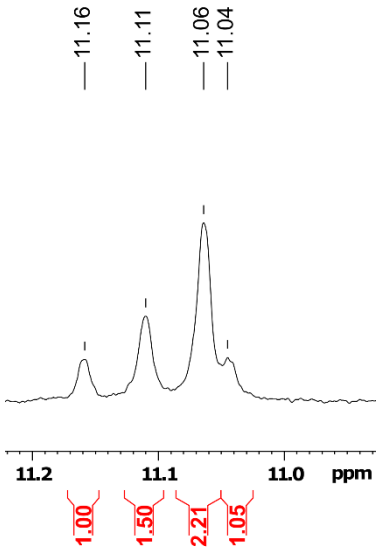


Figure 80. ^{1}H NMR of 13, showing the Trp H$^{\varepsilon 1}$ region at 303 K. Experiment performed at a 1 mM peptide concentration, in DMSO-*d*$_6$, on a 600 MHz spectrometer.

Integration of the peaks showed a relatively even distribution of the five conformations (Table 8). Two of the conformations (with Trp H$^{\varepsilon 1}$ at 11.06 and 11.07 ppm) were identified as *trans* through an ROE between Trp H$^{\alpha}$ and Ala *N*-Me. Although a complete assignment of the backbone was not achieved, an ROE between the methyl group and the H$^{\alpha}$ in (*N*-Me)-Ala was observed, which confirmed that the methyl group was part of the alanine residue. One of the conformations with Trp H$^{\varepsilon 1}$ 11.11 ppm showed an ROE cross peak from the Trp H$^{\alpha}$ to a signal at 4.39 ppm, which would match the chemical shift for an Ala H$^{\alpha}$. This could indicate a *cis* configuration, but this could not be verified as the assignment for the Ala H$^{\alpha}$ could not be confirmed. The remaining two conformations could not be identified as *cis* or *trans*, as no through-space cross peaks were observed.

Table 8. Trp $H^{\varepsilon 1}$ chemical shifts in DMSO-d_6 for the five conformations observed, including the integral, conformation and content for each one.

Trp $H^{\varepsilon 1}$ chemical shift (ppm)	Integral	Conformation	Conformation content
11.04	1.1	Unidentified	18%
11.06	1.1	*Trans*	19%
11.07	1.1	*Trans*	19%
11.11	1.5	Possible *cis*	26%
11.16	1	Unidentified	17%

It was attempted to assign the cysteine residues to determine the origin of the additional conformations, but this could not be achieved as the signals were significantly overlapped in chemical shift and not enough through space cross peaks were present.

In the case of **12**, experiments were run in potassium phosphate buffer at 298 K. A slight distortion of one of the Trp $H^{\varepsilon 1}$ signals was observed, which could indicate the presence of another conformation. Temperature-dependent experiments were run from 278 to 298 K, which at 278 K showed the distorted Trp $H^{\varepsilon 1}$ signals more clearly (Figure 81).

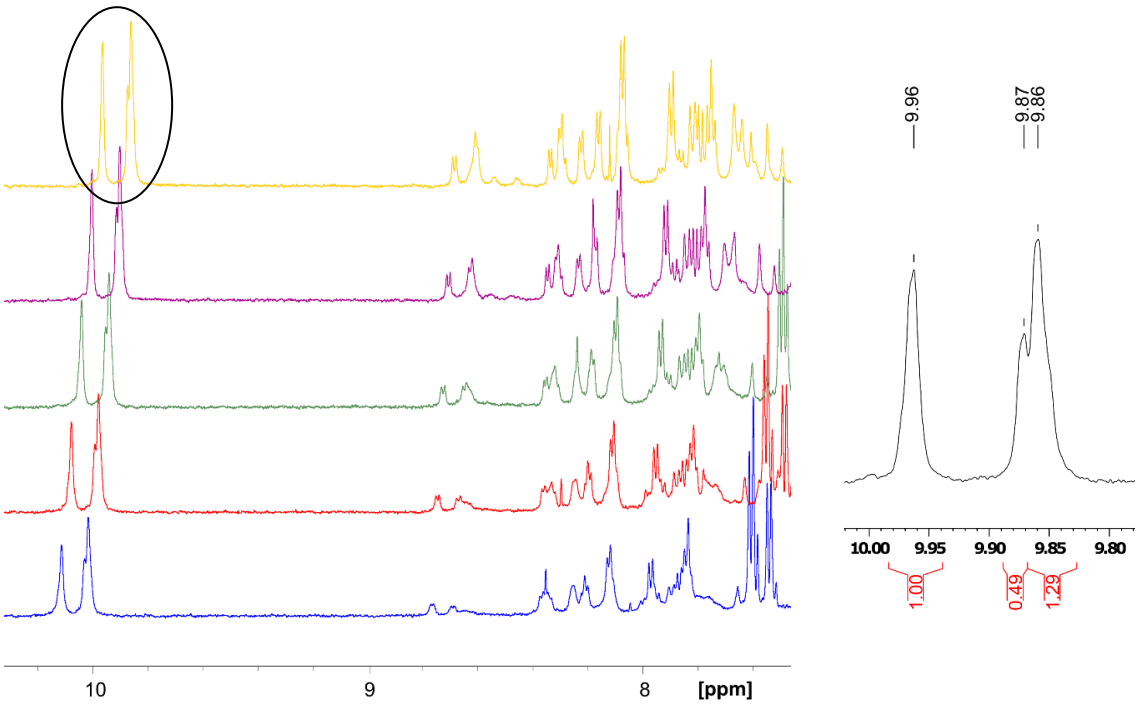


Figure 81. Left: temperature-dependent ^{1}H NMR for 12. From top to bottom: 278, 283, 288, 293 and 298 K. Right: close-up of the Trp H$^{\varepsilon 1}$ signals at 278 K. The experiments were performed at a 1 mM peptide concentration, in potassium phosphate buffer (200 mM, pH 7.4) and 5% D_2O on a 600 MHz spectrometer.

Assignment was performed at 298 K, where three different conformations could be observed based on the Trp H$^{\varepsilon 1}$ signals (Figure 82). The two conformations with Trp H$^{\varepsilon 1}$ at 10.03 and 10.02 ppm were identified as *trans* through an NOE cross peak between Pro H$^{\delta}$ and Trp H$^{\alpha}$.

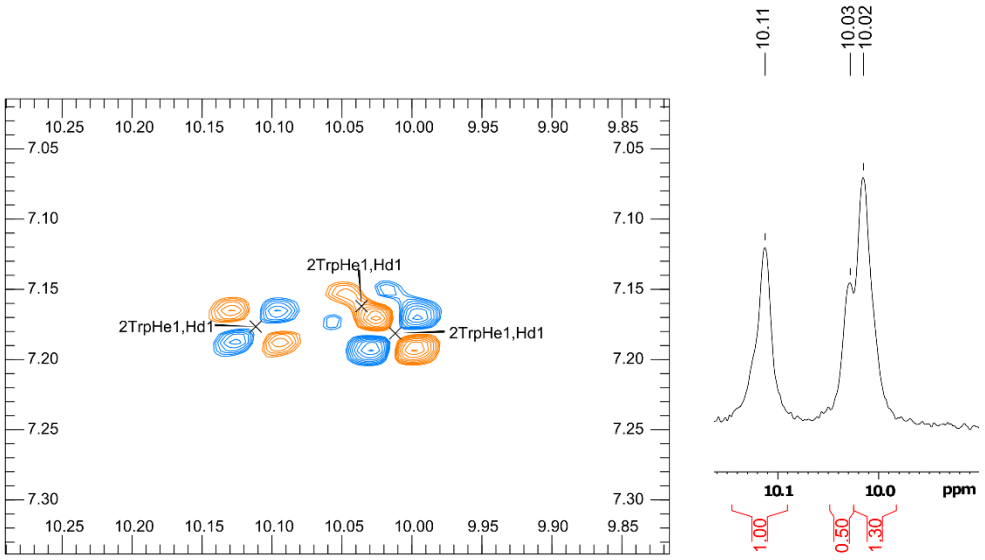


Figure 82. Left: COSY cross peaks between Trp H$^{\varepsilon 1}$ and Trp H$^{\delta 1}$ for the three conformations. The letters ε and δ are represented as e and d, respectively. Right: corresponding ^{1}H NMR signals for the Trp H$^{\varepsilon 1}$ protons. The experiments were performed at a 1 mM peptide concentration, in potassium phosphate buffer (200 mM, pH 7.4) and 5% D_2O, at 298 K on a 600 MHz spectrometer.

The conformation with Trp $H^{\varepsilon 1}$ at 10.11 ppm was identified as *cis* through an NOE between Pro H^{α} and Trp H^{α}. In order to confirm this cross peak, the full assignment of the Pro residue was carried out. Unlike the other two Pro residues, a separate COSY cross peak between the Pro H^{α} and H^{β} was not observed for this conformation. Instead, the H^{α} resonance was overlapped with one of the H^{δ} protons (Figure 83).

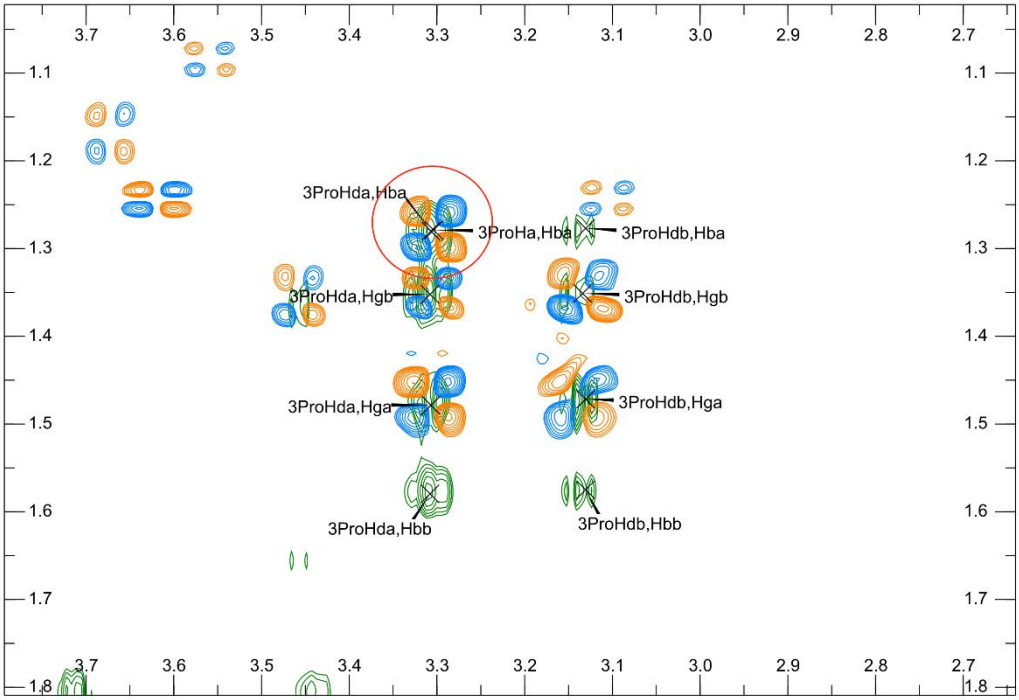


Figure 83. Overlapped COSY (in orange and blue) and TOCSY (green) spectra, in the Pro side chain region for the *cis* configuration. The COSY between Pro H^{α} and H^{β} is circled in red. The letters β, δ and γ are represented as b, d and g.The experiments were performed at a 1 mM peptide concentration, in potassium phosphate buffer (200 mM, pH 7.4) and 5% D_2O, at 298 K on a 600 MHz spectrometer.

The assignment of the cysteine residues was attempted. Correlations between α and β protons were found which allowed the identification of five Cys residues, but due to the backbone NH protons being in exchange with solvent it was not possible to identify which sets of signals corresponded to 4Cys or the *C*-terminal 5Cys, nor which conformation they belonged to. It was therefore not possible to determine the difference between the two *trans* conformations observed. The content of each conformation observed was as follows:

Table 9. Trp $H^{\varepsilon 1}$ chemical shifts for the three conformations observed, including the integral, conformation and content for each one (at 298 K).

Trp $H^{\varepsilon 1}$ chemical shift (ppm)	Integral	Pro conformation	Conformation content
10.11	1	*Cis*	35.7%
10.03	0.5	*Trans*	17.9%
10.02	1.3	*Trans*	46.4%

The *cis* Pro content found for all the peptides studied is shown in Table 10. The occurrence of the *cis* configuration increased by 10% when Pro was substituted by *N*-methylated Ala, which is consistent with the findings from Section 2.5. This suggests that the increased flexibility of the *N*-methylated Ala residue allows the peptide to adopt more geometries in comparison to the more constrained Pro-containing sequence.

Table 10. *Cis/trans* ratios and *cis* content for each peptide at 298 K.

Sequence	*Cis/trans* ratio	*Cis* percentage
Ac-TWPCC-NH_2 reduced cysteines, **10**	0.56	36%
Ac-TWPCC-NH_2 disulfide bridged, **12**	0.56	36%
Ac-TW(*N*-Me)ACC-NH_2 reduced cysteines, **11**	0.85	46%
Ac-TW(*N*-Me)ACC-NH_2 disulfide bridged, **13**	-	-

The two systems studied did not show a two-state switching behaviour upon disulfide formation, but instead generated multiple conformations that could be due to the presence of a *cis* cysteine residue and the disulfide bridge rotamers.

2.8 Conclusions and future work

A double mutant cycle based on the WPQ-containing sequence **1** was used to evaluate the effects of local sequence on Pro conformation. It was found that an aromatic residue preceding proline increased the *cis* Pro conformation content by an order of magnitude compared to alanine, and that positive cooperativity existed between Trp and Gln.

A database search and MD simulations were used to choose the sequences for experimental evaluation and to predict their behaviour. A comparison between the *cis* content estimated by these methods, as well as the values found experimentally by NMR analysis, are presented in Table 11. As it can be observed, the *cis* content for the sequences that do not contain an aromatic residue was more closely predicted by the MD and the database search. In contrast, the predictions for **1** and **3** differed more from the values obtained *via* NMR. In the case of the MD simulations, this could be due to a poor representation of the Trp side chain. The database results could be different because the sequences screened form part of a larger protein structure, whereas the peptides under experimental study lack that additional environment.

Table 11. Comparison of the *cis* Pro content (%) predicted by the database search, the MD simulations and the experimental values obtained through NMR.

	Peptide			
***Cis* Pro %**	**1**	**2**	**3**	**4**
Database	19	8	23	7
MD	31.6	8.2	27.9	8.7
NMR	26	8.3	30	7.4

The *N*-methylated peptide **5** was studied in order to compare the occurrence of the *cis* configuration in a sequence containing (*N*-Me)-Ala with its Pro counterpart **1**, and it was observed that this peptide adopted a higher amount of *cis* (40% for **5** and 26% for **1**). This highlighted the potential of (*N*-Me)-Ala as a substitute for Pro in switching systems based on tertiary amide isomerisation.

The information obtained from the aforementioned studies was used to design a switching system actuated through disulfide formation. Oxidation conditions that yielded the product

in a short reaction time with no dimerisation were screened, which identified a method utilising DPDS as the most suitable. The sequences were studied through NMR, which showed that the peptides containing Trp preceding Pro had ~10 % more *cis* Pro than the peptides containing Ala. The reduced and oxidised peptides did not adopt different Pro conformations and therefore the system did not behave as a conformational switch.

A further design was investigated, where the formation of a vicinal disulfide would function as the actuation method. NMR studies showed that the (*N*-Me)-Ala sequence contained 10% more *cis* configuration in comparison to the Pro-containing system in the reduced cysteines state, and that neither of the peptides behaved as a switching system. Instead, the disulfide-bridged peptides **12** and **13** adopted three and five different conformations respectively, which could be caused by the presence of a *cis C*-terminal cysteine as well as two disulfide bridge rotamers. Future work would involve the assignment of the unidentified conformations in **13**. Changing the pH or the solvent of the NMR sample could reveal more NOE or ROE cross peaks that would allow the identification of a *cis* or *trans N*-methylated Ala. It could also be attempted to change the tumbling regime of the peptide in order to observe more through space correlations.

The cause of the additional conformations adopted by **12** and **13** could be identified by confirming the *cis* or *trans* configuration of the cysteine residues and analysing the disulfide rotamers. The occurrence of a *cis C*-terminal Cys could be confirmed by a ROESY cross peak between the H^{α} protons in the two cysteine residues[260] and the disulfide rotamers could potentially be identified using differences in chemical shifts and exchange cross peaks.[261] A 3D NMR structure of the peptide could be obtained, which would allow the measurement of the χ dihedral angles in the disulfide bridge.

Although a switching system has not been accessed in this work, information has been gained about the backbone of interest. In a future design, (*N*-Me)-Ala could be exploited to access a higher amount of *cis* configuration, in combination with another actuation method. For example, phosphorylation could be employed, which in turn could allow for the transfer of a working system into a biological context. In this setting, a switching peptide could be employed as a biosensor, which for example could allow the detection of kinase activity through the use of a FRET pair.

The backbone of the switch could be improved by drawing inspiration from β-hairpin structures in two ways: firstly, if one or two more amino acids were added to frame the Xaa-Pro-Yaa scaffold, there would be more possibilities to confer stability to the system in the two desired conformations. And secondly, the wide amount of information available about the types of turn that occur in hairpins could be utilised to fine tune the design of a switch that would change from one type of turn to another.

3 β-Turn mimics through hydrazide ligation

The work presented in this chapter was performed in collaboration with Dr Drew Thomson and Dr Albane Neckebroeck. Dr Drew Thomson performed the molecular dynamics simulations. The NMR experiments of the control sequences and their preliminary assignment were performed in collaboration with Dr Albane Neckebroeck.

The design, synthesis and evaluation of **14** in comparison to **TrpZip1** has been published elsewhere.[262]

3.1 Hydrazones and hydrazides in ligation

Hydrazones are attractive moieties for ligation, especially since the development of nucleophilic catalysts like aniline that significantly improve reaction rates. They have been frequently used as ligation anchors in the cyclisation of peptides[197,198] and as peptide bond surrogates,[199] although their use as a stable linkage is limited due to their instability against hydrolysis.[196] This reversibility has been taken advantage of in bioconjugation strategies, as it allows the installation of a linker and its exchange to a new functionality thereafter. For example, Dirksen *et al*.[263] used hydrazone linkers to reversibly tag proteins with biotin (Figure 84). Using this method, the biotin label can be exchanged for a new tag for further analysis and the protein can be recovered capped or not capped for further functionalisation.

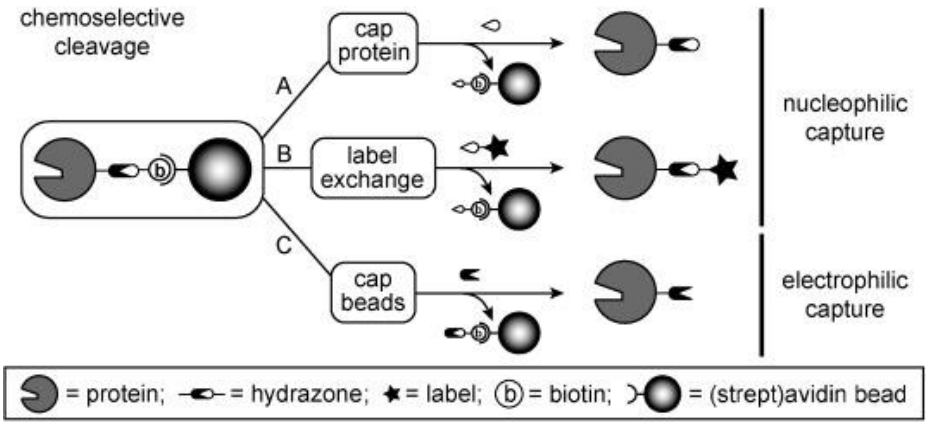


Figure 84. Possible strategies for use with the reversible hydrazone linker designed by Dirksen *et al*. Reproduced with permission.[263]

In addition, this reversibility has been utilised in antibody-drug conjugates to release the active drug below pH 5,[264] to label proteins with fluorescent tags[265] and as a linker to attach a fluorophore onto an enzyme active site. This linker can then be removed through hydrazone exchange to restore enzymatic activity.[266]

Conversely, hydrazides are much more stable against hydrolysis, but their use as a ligation handle in bioconjugation is limited as an additional step is required to reduce the C=N bond. The most widely used application of these moieties is in a variation of the native chemical ligation reaction, which utilises a hydrazide instead of the *C*-terminal thioester (Figure 85). This methodology developed by Fang *et al*.[267,268] allows the use of an unprotected peptide hydrazide, which is converted into a thioester through oxidation followed by reaction with a thiol. This activated species undergoes native chemical ligation with a Cys-containing *N*-terminal peptide, which generates a peptide bond.

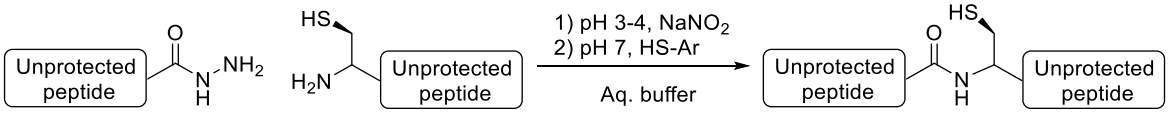


Figure 85. Native chemical ligation of peptide hydrazides.

The stability of the hydrazide linkage has been exploited to install a permanent tag on DNA by Wilkinson *et al*. (Figure 86).[269] In this strategy, DNA is labelled with a bifunctional chemical handle that contains a hydrazone and an azide. This azide is used to introduce a tag by means of an azide-alkyne cycloaddition, which can then be removed by hydrazone exchange. This allows the installation of a new tag by a further hydrazone exchange step, or the permanent tagging of DNA with a hydrazide linkage through the reaction with an *N*-hydroxysuccinimide ester.

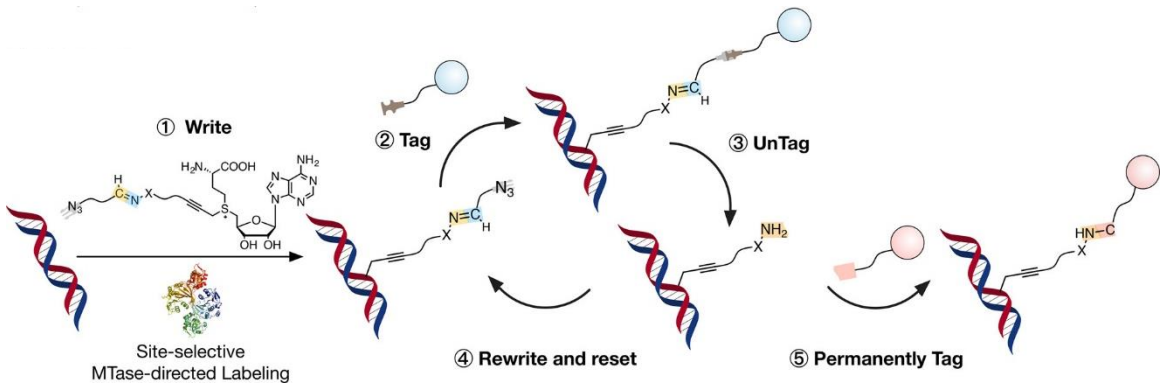


Figure 86. DNA tagging and untagging method developed by Wilkinson *et al*. Reproduced with permission from ref.[269]

3.2 Hydrazone and hydrazide-containing β-turn mimics

Hydrazide and hydrazone linkages have been used as components of mimics of protein structure in the literature, but the vast majority of examples are used as hydrogen bond surrogates[142] within α-helices and as peptide bond surrogates[199,270] within short peptide backbones. An example of the latter is the work by Zerkout *et al.*,[271] who substituted Pro for its hydrazino analogue in tripeptides of the sequence Boc-Gly-Pro-Gly-NH^{i}Pr and compared the in-solution and crystal structures of both analogues. They found that the peptide containing Pro adopted a type II β-turn, whereas the analogue containing hydrazino-Pro adopted a different folded conformation. In the analogue with hydrazino-Pro, a bifurcated hydrogen bond took place (Figure 87) between the carbonyl in 1Gly, the hydrazino-Pro N$^{\alpha}$ and the 3Gly NH. The authors suggested that this 'hydrazino fold' acted as a β-turn mimic.

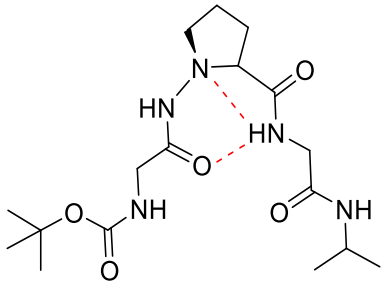


Figure 87. Hydrazino-Pro-containing tripeptide studied by Zerkout *et al.*[271] The bifurcated hydrogen bond is depicted with a red dashed line.

Lubell *et al.*[170,272] have made use of the hydrazide-containing residues Aza, Aid and Nai (Figure 88) as β-turn mimics in the peptide *N*-4-bromobenzoyl-Aid-Phe-*N*'-*iso*-propylamide. Although a hydrazide linkage is present in the unnatural unit, ligation chemistry is not used to install that moiety – instead, an acylation step is required.

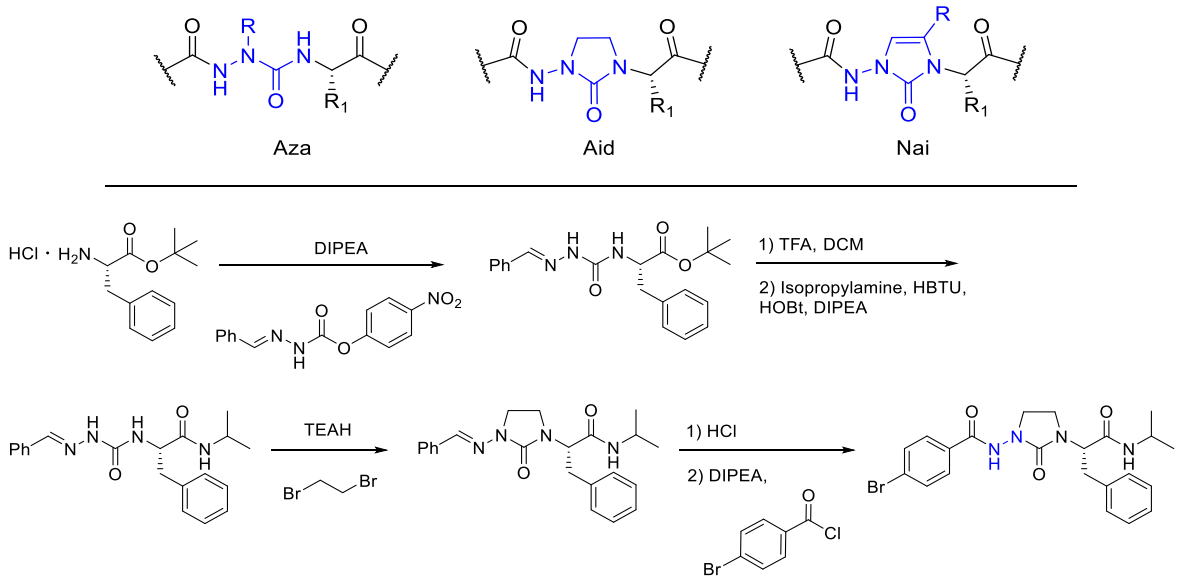


Figure 88. Top: from left to right: Aza, Aid and Nai peptidomimetics. Bottom: synthesis of *N*-4-bromobenzoyl-Aid-Phe-*N'*-*iso*-propylamide by Lubell *et al*.[170]

The only example in the literature of a hydrazide turn generated through a ligation step is by Vanderesse *et al*. (see Section 1.6.2),[146,273] who used this linkage as well as an oxime, amidoxy and a reduced amidoxy moiety as amide surrogates to induce γ-turns in dipeptides and hexapeptides (Figure 89).

Ala-Ala-Pro-*link*-Val-Ala-Ala-NH^{i}Pr

Ala-Ala-Pro-Val-*link*-Ala-Ala-NH^{i}Pr

[CH=N-O] Oxime

[CH_2-NH-O] Reduced amidoxy

[CO-NH-NH] Hydrazide

[CO-NH-O] Amidoxy

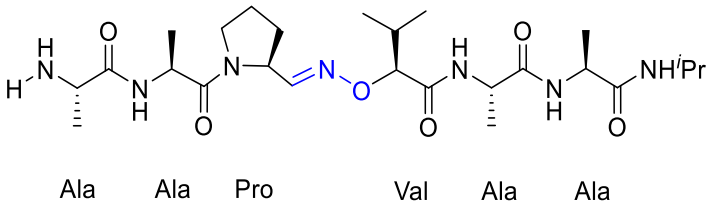


Figure 89. Peptide sequences (left) used by Vanderesse *et al*.[146,273] to test the different peptide bond surrogates (right), and chemical structure of one of the hexapeptides.

3.3 Chapter aims

The modification of β-turns has received considerable attention in the literature (as has been described in Section 1.6.4), as mimicking this secondary structure provides a useful tool in a wide variety of areas, from understanding protein folding to targeting diseases. But although numerous β-turn mimics have been previously designed, these are either accessed through multistep organic syntheses (like Hot=tap[174]) or they are based on simple building blocks which lack structural preorganisation (like ornithine[163]). This significantly limits their possibilities for application in peptide and protein science.

The aim of this work was to design new β-turn mimics that would provide the necessary preorganisation without the need for a multistep organic synthesis. This would be achieved through the use of ligation chemistry as a means to install the mimic structure. To this end, two peptides with unnatural modifications would be used as the ligation partners and the β-turn mimic would be generated *in situ*, utilising the ligation junction functionality.

The aims of this chapter were as follows:

- Design of the β-turn mimic scaffolds.
- Synthesis of the β-turn mimics within the β-hairpin backbone chosen as well as in small peptides.
- Development of efficient ligation conditions.
- Synthesis of the chosen model system, a tryptophan zipper (TrpZip).
- Evaluation of the β-turn mimics using CD and NMR to compare their conformation and thermal stability with the control systems.

3.4 Choice of model system and design

The β-hairpin **TrpZip1**, designed by Cochran *et al*.,[80] has been used as the model system for these studies (Figure 90). **TrpZip1** is a designed 12-residue tryptophan zipper (for the complete sequence see Section 7.12) that folds into a soluble, monomeric hairpin structure in solution, containing a type II' β-turn. The 3D structure of this peptide has been extensively

characterised through NMR, and data on the thermodynamics of its folding are available. In addition, it has been shown that the modification of the β-turn in **TrpZip1** can alter the stability of the peptide fold. Therefore, this peptide was an ideal base structure to use for the investigation of the conformational and thermodynamic effects of modifying the β-turn. A further control peptide was used in these studies, namely **TrpZipGG**, where the turn residues *i+1* and *i+2* were substituted by Gly-Gly in order to compare to an achiral turn (Figure 90).

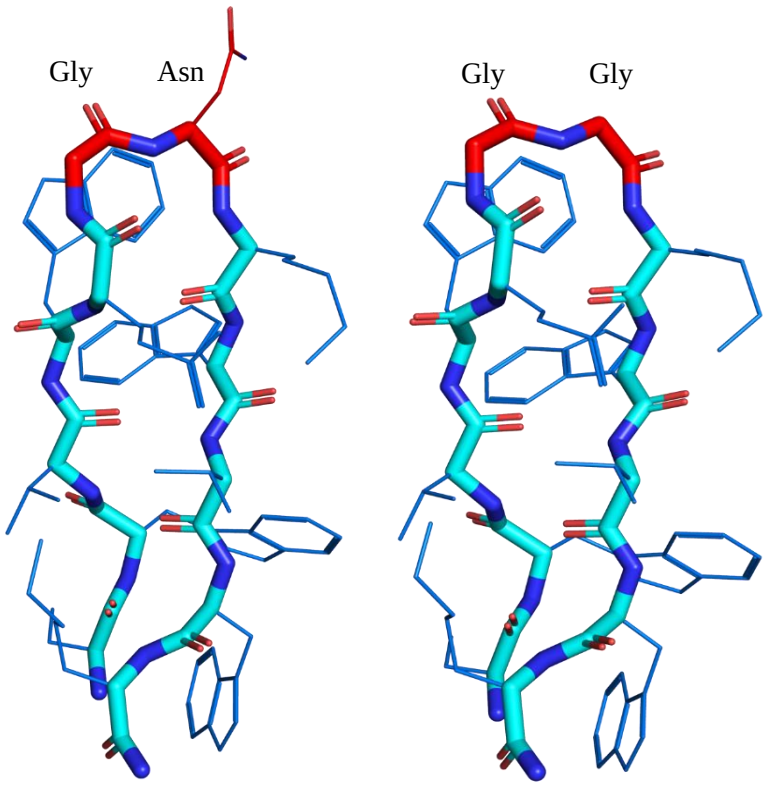
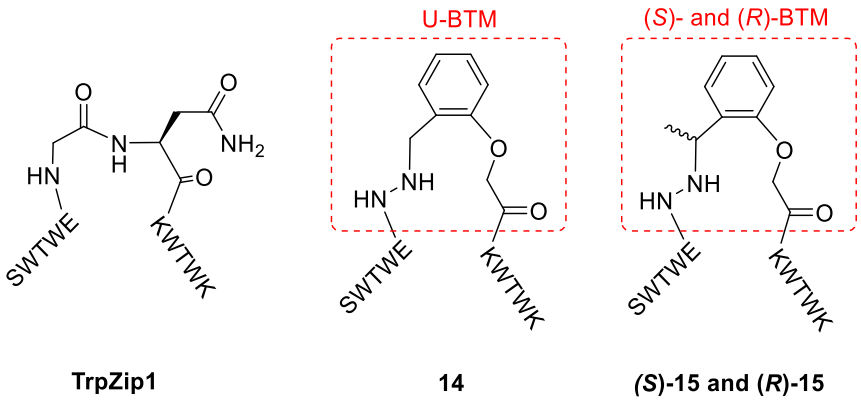


Figure 90. NMR structure of TrpZip1 (left, PDB 1LE0) and TrpZipGG (right, generated through PyMol mutation and energy minimised). The *i+1* and *i+2* residues in the turn are represented in red.

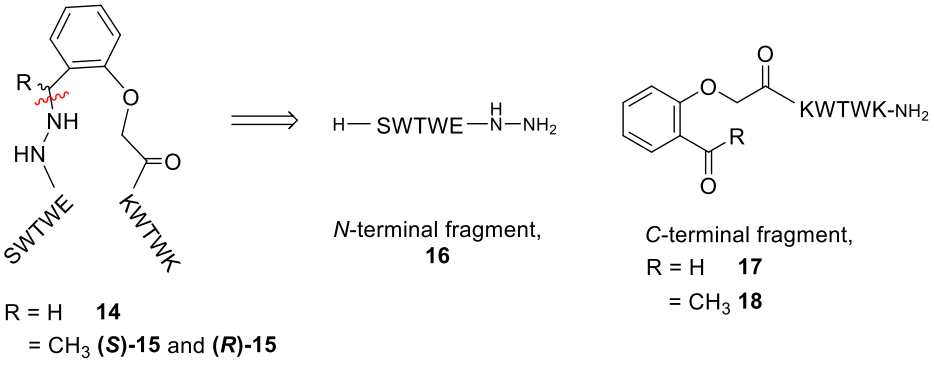
The β-turn mimics in study were to be installed in place of the *i+1* and *i+2* residues in the turn through the chemoselective ligation of two peptide fragments. This would take advantage of the functionality in the ligation junction to generate the β-turn in the hairpin. A hydrazide linkage was envisioned, which would be generated by a reductive amination step between two peptide partners. The designed β-turn scaffolds consisted of an *ortho*-substituted aromatic ring, which would favour the formation of the hydrogen bond between the positions equivalent to *i* and *i+3* in the native peptide. Three β-turn mimics would be studied, one with a methylene unit in the benzylic position (unsubstituted β-turn mimic, U-

BTM) and the two isomers arising from the presence of a methyl substituent on the benzylic position ((*S*)-BTM and (*R*)-BTM) (Scheme 14).



Scheme 14. TrpZip1 control and hydrazide-linked β-turn mimics studied.

The peptide substrates for the ligation would consist of an *N*-terminal fragment with a hydrazide moiety at the *C*-terminus (**16**) and a *C*-terminal fragment modified with the carbonyl-containing aromatic unit in the *N*-terminus (Scheme 15). Two options were to be explored for the *C*-terminal peptide: one with an aldehyde in the benzylic position (**17**) and another containing a ketone (**18**).



Scheme 15. *N*-terminal and *C*-terminal peptide substrates for the ligation.

This strategy would provide a straightforward manner to install the mimic, which would be compatible with unprotected peptide substrates. This property would be novel within the literature, as other mimics reported involve more complex synthetic routes, and it could be useful in the study of β-hairpin structures and protein folding more generally.

In addition to the peptide conjugates, tripeptide systems containing the β-turn mimic scaffolds flanked by two alanine residues would also be studied. These systems would be

used to aid the NMR assignment of the β-hairpin conjugates, and their conformation would also be compared to the full peptide structures to ascertain if they fold in solution.

3.5 Molecular dynamics simulations

Molecular dynamics simulations were run by Dr Drew Thomson on both the small peptide systems and the peptide conjugates in order to evaluate the potential of the turn units to function as β-turn mimics. For this, position restrained energy minimisation was performed for each peptide, followed by an MD simulation over 100 ns.

The models obtained for the peptide conjugates (Figure 91) showed a similar secondary structure to **TrpZip1**, exhibiting a hydrogen-bonded β-hairpin conformation with the same aromatic stacking pattern and polar face. In the case of **(*S*)-15** and **(*R*)-15**, both isomers adopted similar conformations.

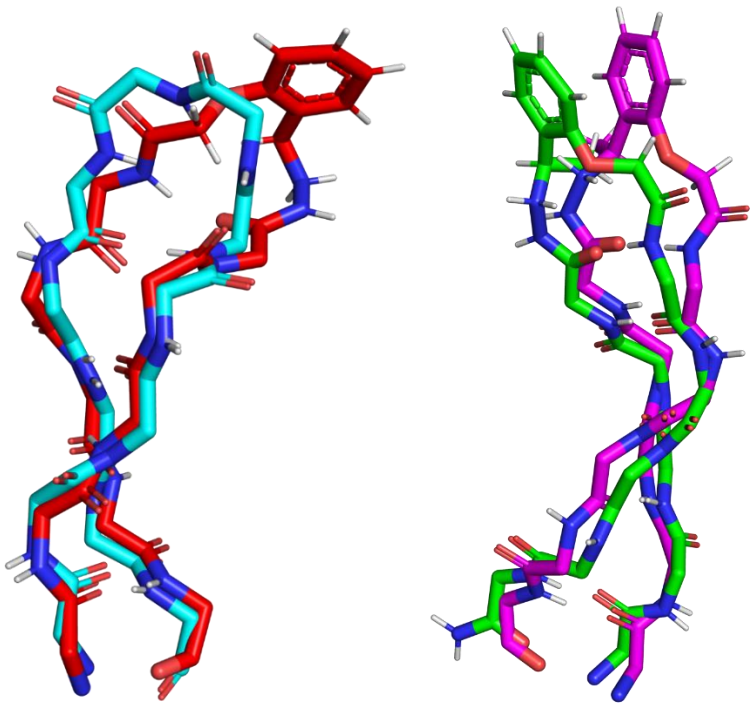


Figure 91. Models obtained from the MD simulations, showing only the peptide backbones. Left: 14 (red) overlayed with TrpZip1 (blue), and right: (*S*)-15 (magenta) and (*R*)-15 (green).

Importantly, the β-turn-like hydrogen bond between the residues preceding and following the β-turn mimic was maintained in all peptide conjugates (Figure 92). The distance between the oxygen in the carbonyl of the residue preceding U-BTM and the proton in the amide NH of the residue following U-BTM was calculated along the MD trajectory for **14**, in order to compare it to the same distance between *i* and *i+3* in **TrpZip1**.

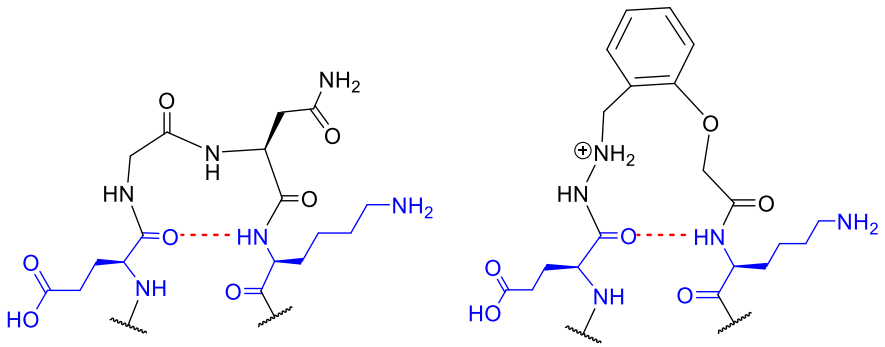
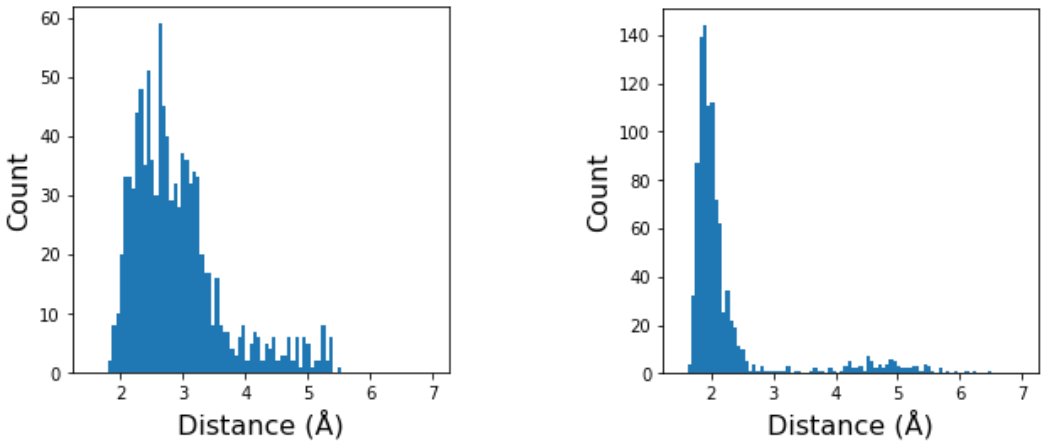


Figure 92. Distances between the *i* and *i+3* residues (in blue) measured using MD (in red) in TrpZip1 (left) and 14 (right).

The histograms of the distances obtained are depicted in Figure 93. The distance between the atoms was smaller than 3 Å for 90% of frames during the trajectory of **14**, compared to 63% of frames for **TrpZip1**. This indicated that the hydrogen bond in **14** was more stable during the simulation in comparison to the hydrogen bond in the chosen control. The same analysis was performed for **(*S*)-15** and **(*R*)-15**, which showed an analogous behaviour to **14** (see Sections 7.3.2.1 and 7.3.2.2). All in all, these simulations strongly supported the potential of U-BTM, (*S*)-BTM and (*R*)-BTM to function as β-turn mimics.



Mean = 2.92

SD = 0.74

Mean = 2.26

SD =0.85

Figure 93. Histograms and corresponding distances (in Å) between the oxygen in the C=O(*i*) and the proton in NH(*i+3*) for TrpZip1 (left) and 14 (right).

Models were also generated for the tripeptide systems (Figure 94), which consisted of the sequence Ac-Ala-*turn*-Ala-NH_2, where *turn* = U-BTM (**19**), (*S*)-BTM or (*R*)-BTM (**20a**, **20b**). These models did not show a folded conformation during the majority of the simulation, but also adopted open conformations.

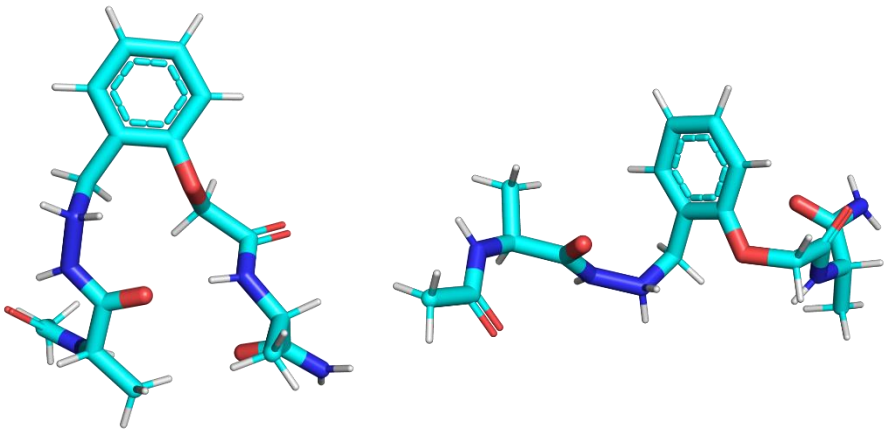
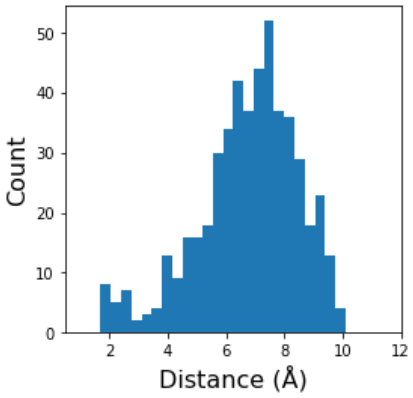


Figure 94. Selected frames from the MD simulation of 19, showing a folded and open conformation.

The distance between the atoms involved in the β-turn-like hydrogen bond was measured for the three analogues in order to ascertain whether the β-turn would be present in systems with a very short β-hairpin sequence in comparison to the full-scale conjugates. As it can be observed in Figure 95, the distances were larger than what would be expected for a hydrogen bond, and therefore indicated that no persistent structure was taking place for any of the analogues. The lack of structure would be confirmed through experimental models, which would also be used to explore ligation conditions and to facilitate the NMR assignment of the full-scale conjugates.

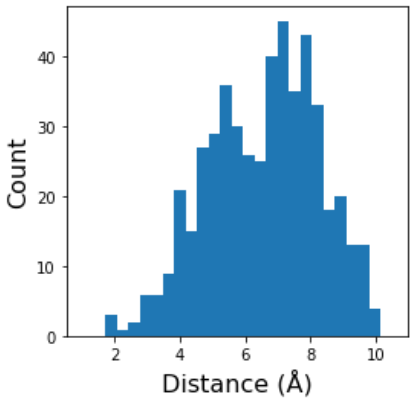
A) Analogue **19**



Mean = 6.76

SD = 1.72

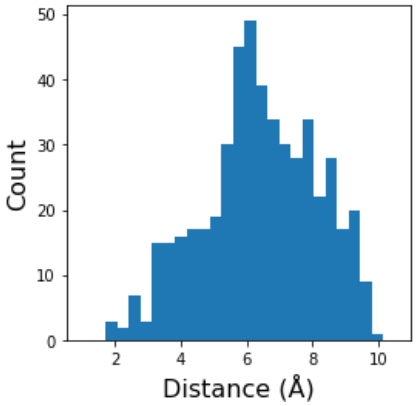
B) Analogue **20**, (*S*)-diastereomer



Mean = 6.56

SD = 1.70

C) Analogue **20**, (*R*)-diastereomer



Mean = 6.41

SD = 1.73

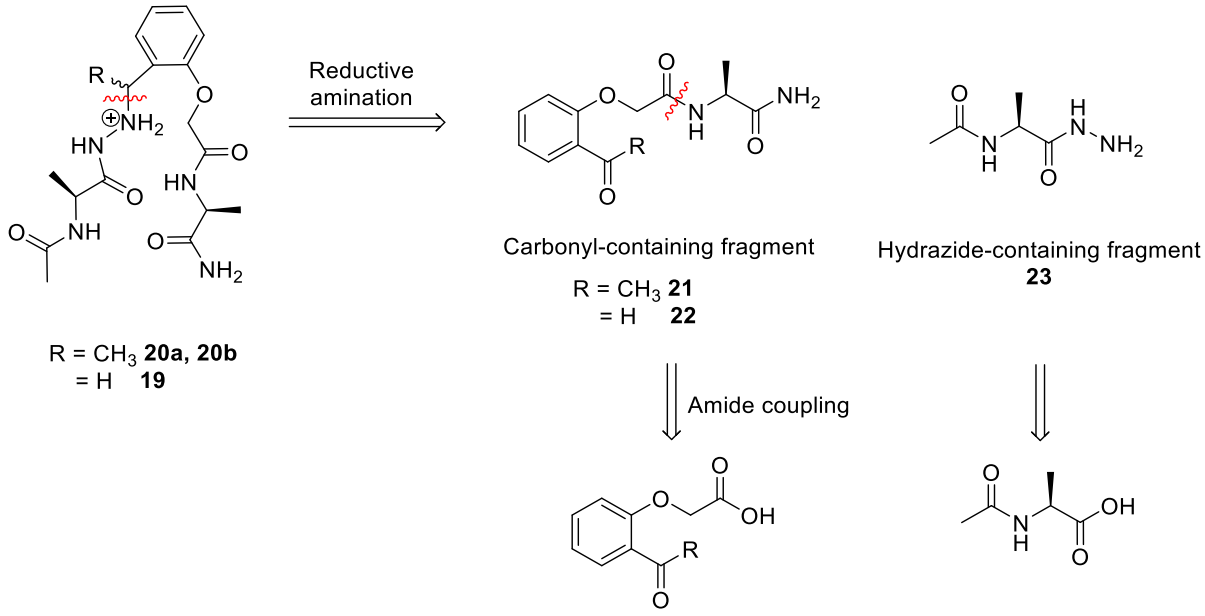
Figure 95. Histograms and corresponding distances between the C=O and NH extracted from the MD simulations for the β-turn-like hydrogen bond in the small peptide systems. Distances are presented in Å.

3.6 Tripeptide systems

3.6.1 Synthesis of 19, 20a and 20b

The turn mimics flanked by two alanine residues were to be synthesised in order to compare their conformation to the full peptide structures. The synthetic route proposed consisted of a reductive amination between the carbonyl and the hydrazide-containing fragments, equivalent to the one to be performed with the peptide conjugates. The carbonyl-containing fragment (**21** and **22**) was to be accessed through a coupling step starting from the commercially available 2-formylphenoxyacetic acid and 2-acetylphenoxyacetic acid, while

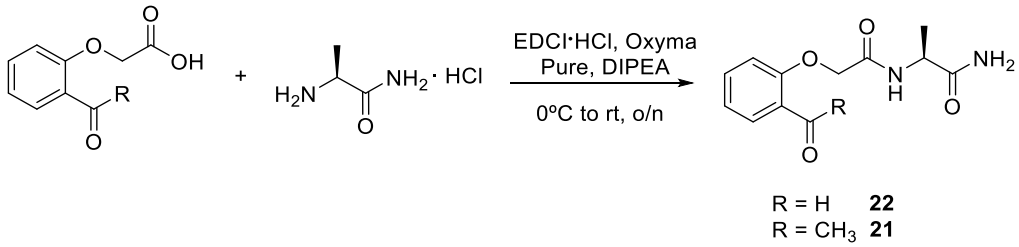
the hydrazide-containing fragment **23** was to be generated through multiple steps from acetyl alanine (Scheme 16).



Scheme 16. Retrosynthetic analysis for the generation of the turn mimic.

3.6.1.1 Synthesis of the carbonyl fragments 21 and 22

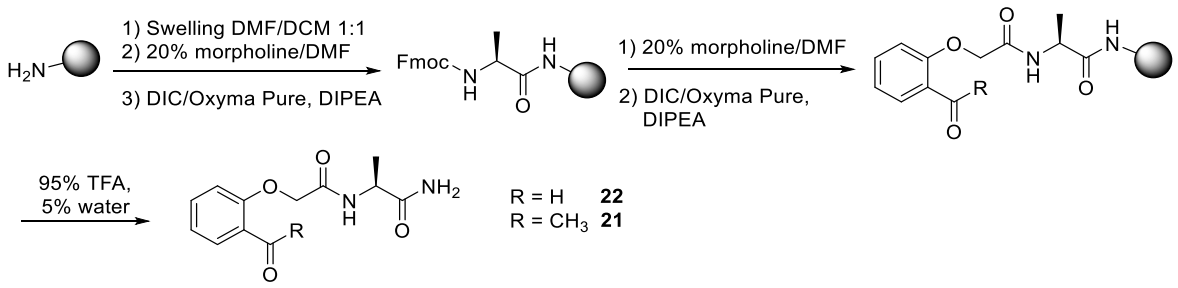
Solution-phase conditions were tested using EDCI·HCl (1.5 equiv.), Oxyma Pure and DIPEA (Scheme 17), which afforded **21** in a 30% yield and **22** in a 2% yield. When removing the base **21** and **22** were obtained in a 31% and 6% yield, respectively.



Scheme 17. In-solution coupling conditions tested for 21 and 22.

In search for better yielding conditions, DIC was tried as the coupling agent, as well as less equivalents of base, but no improvement was observed. Conditions on-resin were to be tested next, in order to find a protocol that could be applied to the larger peptide systems (Scheme 18). For this, Fmoc-Rink Amide AM Resin (0.51 mmol/g loading) was used. Fmoc-Ala was coupled using standard coupling conditions with DIC/Oxyma Pure at room temperature for 2.5 h, and after deprotection 2-acetylphenoxyacetic acid or 2-formylphenoxyacetic acid were coupled using the same conditions. Cleavage from the resin was performed using only water

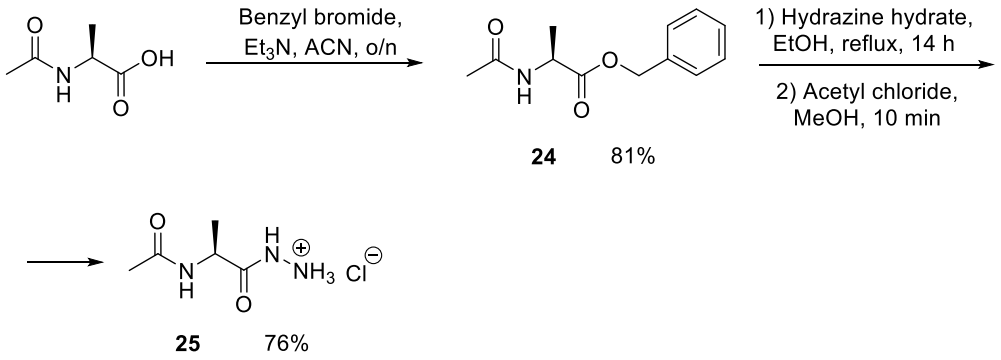
as the scavenger to prevent the reduction of the aldehyde by TIPS. This synthetic strategy afforded **21** in a 67% overall yield and **22** in a 45% overall yield. Interestingly, the yields significantly improved in comparison to the solution-phase synthesis – this could be due to the change of coupling agent and the use of more equivalents in the on-resin procedure (4.5 equiv. of DIC).



Scheme 18. Solid phase strategy for the synthesis of 21 and 22.

3.6.1.2 Synthesis of the hydrazide fragment, 23

The initial strategy proposed for the synthesis of **23** started with an S_N2 reaction between acetyl alanine and benzyl bromide, which furnished the pure product after work-up in an 81% yield (Scheme 19). With **24** in hand, the hydrazide was generated through a reflux with hydrazine hydrate in ethanol. It was attempted to purify **23** through column chromatography using literature conditions,[274] using DCM/MeOH 80:2, but the product could not be recovered from the column. Conditions to induce the precipitation of **23** were then sought: for this, acetyl chloride and methanol were used, followed by trituration in cold acetonitrile, which afforded the HCl salt **25** in a 76% yield.

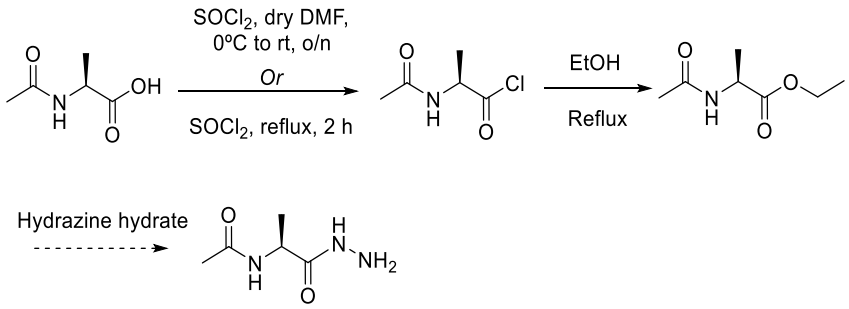


Scheme 19. Initial synthetic route towards 25.

During the characterisation step of this compound, it was noticed that the mass weighed for NMR analysis did not correspond with the very weak signals observed in the spectrum. However, no additional peaks were observed in the ^{1}H NMR nor ^{13}C NMR spectra. It was

proposed that, although concentration under vacuum was performed in between steps, residual hydrazine·HCl could be co-precipitating with the desired product.

A second synthetic plan was then pursued (Scheme 20), which was a one-pot strategy consisting of the formation of the acyl chloride with $SOCl_2$, followed by a reflux in ethanol and finally a nucleophilic substitution with hydrazine to yield **23**. The expected benefit of this strategy was a facile purification after the second step, which would provide pure product after the hydrazide formation.

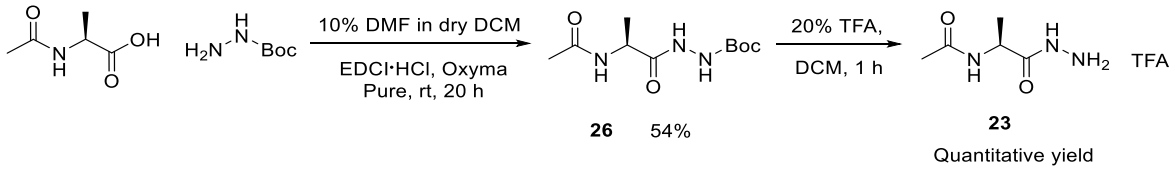


Scheme 20. One-pot synthetic strategy towards 23.

The conditions tested involved performing the first step in dry conditions and adding 1.2 equivalents of $SOCl_2$ to the reaction mixture at 0 °C. TLC monitoring after the reaction was left overnight at room temperature showed product formation, although there was leftover starting material. The reaction crude was then refluxed in EtOH immediately, as the acyl chloride could be unstable. NMR analysis of the reaction crude after this step showed product formation, although the spectrum was complex and a considerable amount of material was lost after work-up (30 mg obtained from 300 mg of starting material). A possible explanation for such a loss in material could be that the first step was performed in DMF, which could potentially react with thionyl chloride.

The second attempt was performing the first step neat and under reflux, while keeping the same conditions for the second step – this showed no product mass by LCMS analysis after the ester formation.

The third and final strategy considered to access the modified alanine fragment was to perform a coupling between acetyl alanine and *tert*-butyl carbazate (Scheme 21). Using conditions by Lascano *et al.*,[275] **26** was obtained in a 54% yield. Subsequent deprotection of the hydrazide using 20% TFA in DCM afforded **23** in quantitative yield. Due to the facile purification of this compound, this was the strategy chosen for synthesis.



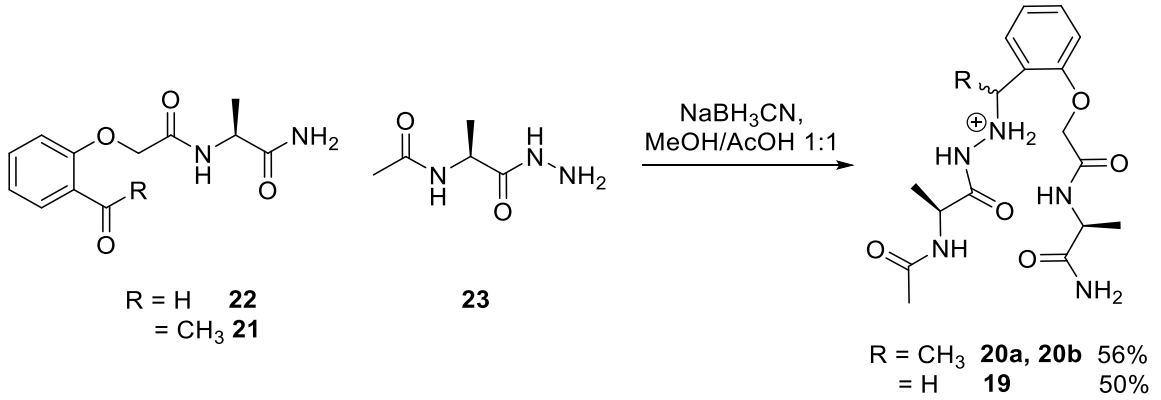
26 54%

23
Quantitative yield

Scheme 21. Final synthetic strategy towards 23.

3.6.1.3 Reductive amination

The reductive amination between **23** and **21** or **22** was performed in MeOH/AcOH 1:1 and using an excess of sodium cyanoborohydride as the reducing agent (Scheme 22). This afforded compounds **20a** and **20b**, which were named as such according to the order in which they are eluted from the column during RP-HPLC analysis.



R = H **22**
= CH_3 **21**

23

R = CH_3 **20a, 20b** 56%
= H **19** 50%

Scheme 22. Reductive amination towards the Ala-flanked turn mimics.

HPLC monitoring was performed for these reactions: for **19** (Figure 96), it was observed that the reaction reached completion in 30 min. Interestingly, a broad peak (at 10.5 min) was observed at the start of the reaction, which was consumed when the reducing agent was added. This was identified as the hydrazone-linked compound, which could be pre-formed when mixing the starting materials. The broadening of the peak may be due to exchange between the *Z*- and *E*-hydrazone configurations. A side product was generated during this reaction (at 12 min), which could not be identified.

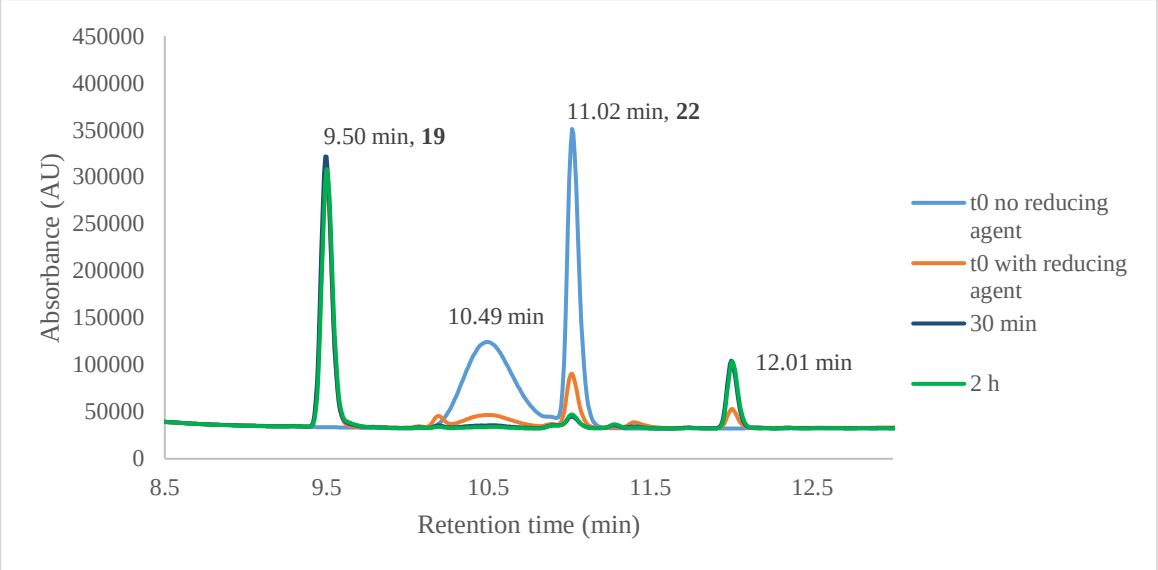


Figure 96. HPLC monitoring of the reductive amination reaction towards 19.

HPLC monitoring of the reductive amination towards **20a** and **20b** was also performed (Figure 97), which showed that the reaction was slower in this case, reaching completion after 3 h.

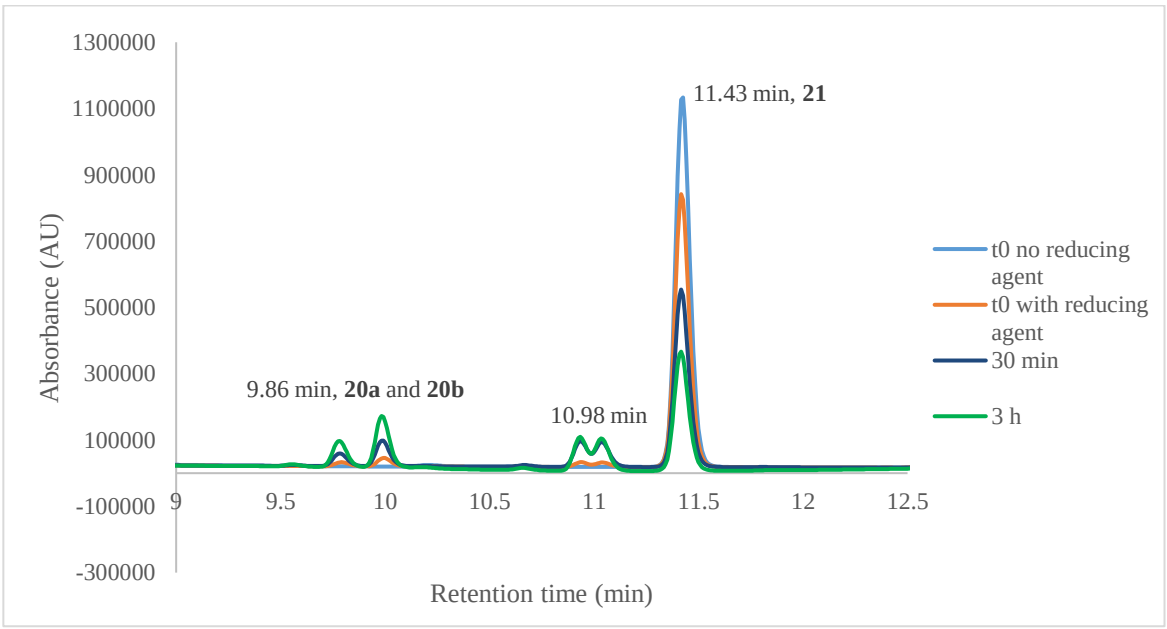
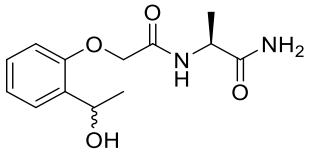


Figure 97. HPLC monitoring of the reductive amination reaction towards 20a and 20b.

Two by-products were also generated during this step (at a retention time of 11 min), which were identified as the product of the reduction of the ketone moiety in **21** to an alcohol (Scheme 23).

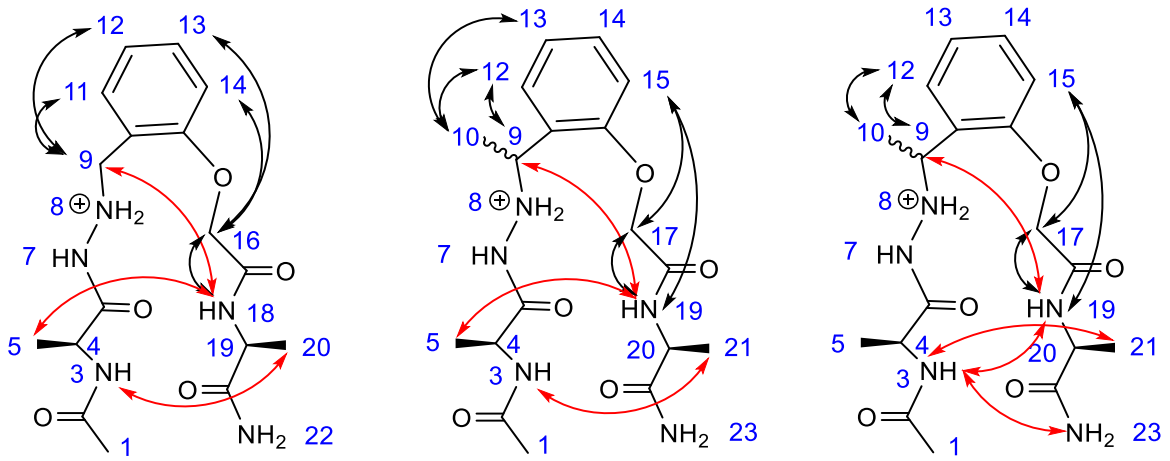


Scheme 23. By-product generated during the reductive amination reaction towards 20a and 20b.

3.6.2 NMR analysis

The Ala-flanked turn mimics were evaluated through NMR in AcOH-d_4 buffer (10 mM, pH 5.5), in 5% D_2O and using 3-(trimethylsilyl)-1-propanesulfonic acid-d_6 sodium salt (DSS) as the internal standard, at 288 K and at concentrations of 14 mM (**19**), 5 mM (**20a**) and 12 mM (**20b**).

Complete assignment of all signals was achieved using COSY, TOCSY, NOESY as well as ROESY through-space experiments, the hydrazide (number 7 in Figure 98) and hydrazinium (8) protons being the only ones to not appear on the spectrum due to exchange with the solvent. The most relevant through-space correlations observed are shown in Figure 98.



19 **20a** **20b**

Figure 98. Key through-space NOE and ROE cross peaks found for 19 (left), 20a (centre) and 20b (right). The cross peaks used for structure confirmation are shown in black and other through-space interactions in red.

The connectivity of the structure of **19** was confirmed through the cross peaks between the CH_2 number 9 and the aromatic ring, as well as between the aromatic ring and the CH_2 16. The equivalent was observed for **20a** and **20b**, as well as a cross peak between the CH_3 number 10 and the aromatic ring.

Correlations were observed that indicate a stable spatial conformation: for **19**, an NOE was present between CH_2 9 and NH 18, equivalent to CH 9 and NH 19 for **20a** and **20b**. In addition, for **19**, cross peaks between 3 and 20 and between 5 and 18 were observed, which were also present in the other two analogues. This indicated proximity between the two Ala residues, confirming that a somewhat folded conformation was taking place. Interestingly, no cross peaks involving the H^{α} were observed, possibly pointing towards a β-hairpin-like conformation where the Ala H^{β} would be closer to the NH in the opposite strand than the H^{α}, the latter pointing out towards the sides of the structure. If this were the case, a hydrogen bond could be taking place between the two Ala residues in the compound. Indeed, a cross peak was observed for **20b** between NH 3 and NH 19, which was diagnostic for a hydrogen bond taking place between the carbonyl 6 and the NH 19. This cross peak was also observed for **20a**, although very weak in comparison. All in all, these NOE and ROE correlations suggested that although these are dynamic systems, the Ala-flanked turn unit was adopting a turn conformation in solution some of the time.

An attempt was made to identify the two diastereomers in **20a** and **20b**, but this could not be achieved as the cross peaks from CH 9 and CH_3 10 to the aromatic ring were equivalent for both compounds, and no further through space interactions involving those protons were present. The purpose of these sequences was to determine whether they adopted folded conformations in solution and to facilitate the NMR assignment of the full-sized peptide conjugates. The first aim was achieved, with all three analogues adopting stable conformations and the *i* to *i+3* like hydrogen bond being present in **20b** and possibly **20a**. The lack of through space NOE correlations with the CH 9 and CH_3 10 in **20a** and **20b** meant that not much additional information could be obtained to use in the full-sized peptide conjugates, and therefore the assignment of the diastereomers was not pursued further.

3.7 Synthesis of the control peptides

Two control sequences were to be synthesised, namely **TrpZip1** and **TrpZipGG**, the latter containing Gly-Gly instead of Gly-Asn in the *i+1* and *i+2* positions of the β-turn. **TrpZipGG** was selected as an additional control in order to compare the behaviour of the peptide conjugates to an achiral turn. The synthesis of the two control sequences was performed in collaboration with Dr Albane Neckebroeck.

TrpZip1 and **TrpZipGG** (Figure 99) were synthesised through standard microwave-assisted SPPS, using the Fmoc/^{t}Bu strategy. DIC/Oxyma Pure activation was employed on a 0.1 mmol scale on TentaGel S RAM resin (loading 0.24 mmol/g). Cleavage from the resin was performed using 95% TFA, 2.5% H_2O and 2.5% TIPS for 2 h. After purification through RP-HPLC, **TrpZip1** was obtained in a 14% yield and **TrpZipGG** in an 18% yield.

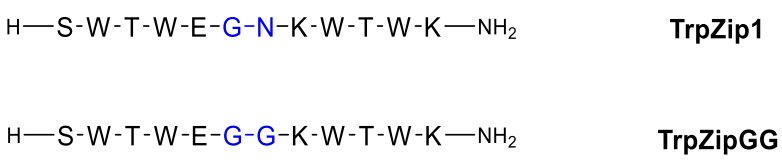


Figure 99. Sequences of the control peptides, TrpZip1 and TrpZipGG.

3.8 Synthesis of the peptide conjugates 14, (*S*)-15 and (*R*)-15

The three target peptide conjugates had the same sequence as **TrpZip1**, where the residues Gly-Asn in the β-turn were to be substituted by the β-turn mimic units (Figure 100).

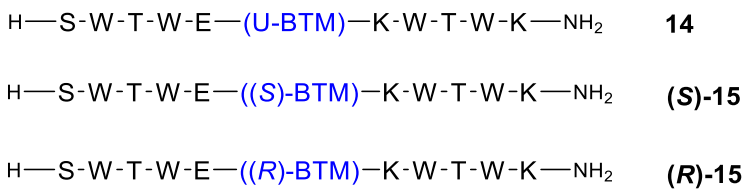
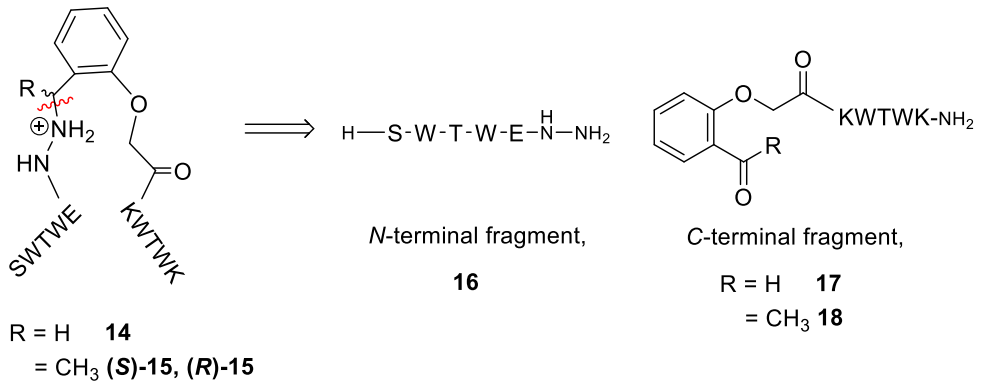


Figure 100. Sequences of the target peptide conjugates.

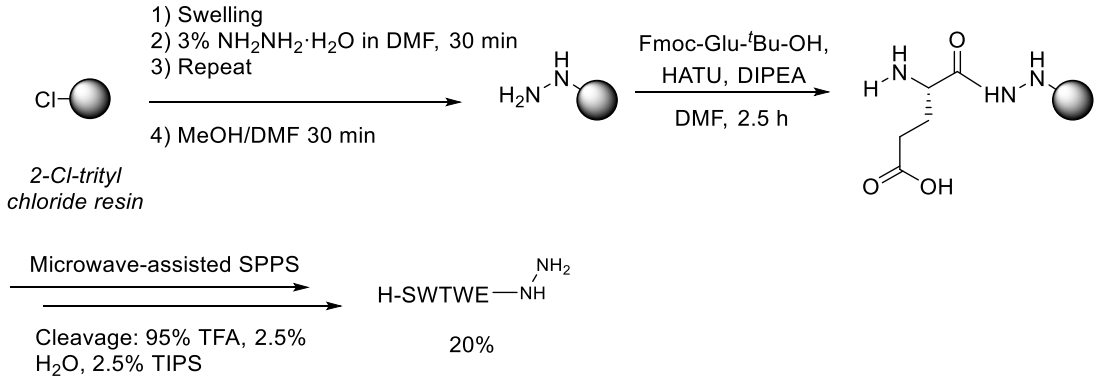
These sequences were to be accessed through a reductive amination between two peptide substrates, which would generate a hydrazide linkage (Scheme 24). The *N*-terminal fragment was modified with a hydrazide moiety in the *C*-terminus (**16**) and the *C*-terminal fragment contained a non-natural aromatic unit in the *N*-terminus, either with an aldehyde (**17**) or a ketone moiety (**18**).



Scheme 24. Disconnection to generate the target peptide conjugates.

3.8.1 Synthesis of the *N*-terminal and *C*-terminal fragments

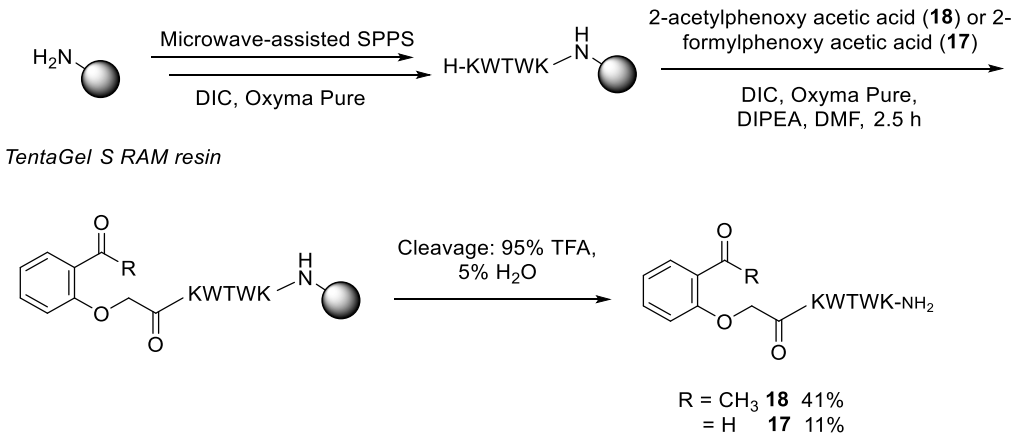
The *N*-terminal fragment was synthesised on 2-Cl-trityl chloride resin (0.74 mmol/g loading) on a 0.1 mmol scale (Scheme 25). The hydrazide modification was introduced by treating the resin with a solution of 3% hydrazine hydrate in DMF for 30 min, and repeating that process. The capping of any unreacted resin was then performed, using 10% MeOH in DMF for 30 min. Subsequently, the first amino acid in the sequence was coupled with HATU for 2.5 h at room temperature and the rest of the sequence was synthesised through standard microwave-assisted SPPS using DIC/Oxyma Pure activation. Cleavage from the resin was performed using 95% TFA/2.5% water/2.5% TIPS. This afforded **16** in a 20% overall yield after purification.



Scheme 25. Synthetic route to access 16.

The two *C*-terminal fragments were synthesised on a TentaGel S RAM resin (0.24 mmol/g loading) on a 0.05 mmol scale (Scheme 26). The synthesis was performed through standard microwave-assisted SPPS using DIC/Oxyma Pure activation until the final Lys deprotection step. At that stage, the unnatural unit (2-acetylphenoxyacetic acid for **18** and 2-formylphenoxyacetic acid for **17**) was coupled using DIC/Oxyma Pure at room temperature for 2.5 h. Cleavage from the resin was performed without the use of TIPS to prevent the

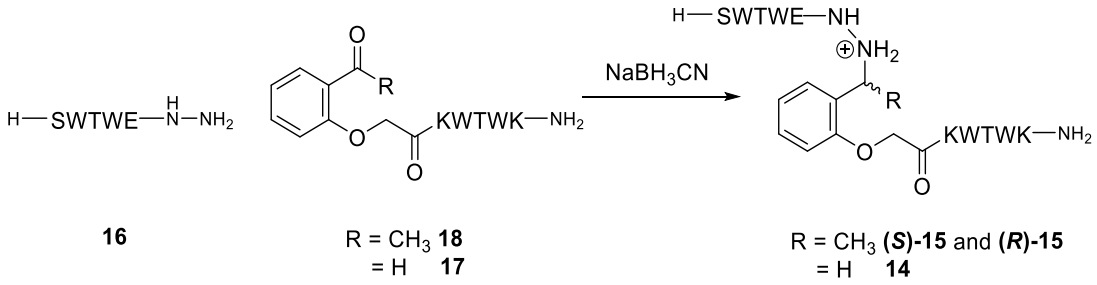
reduction of the aldehyde and ketone moieties. This afforded **17** in an 11% yield and **18** in a 41% yield after purification *via* RP-HPLC.



Scheme 26. Synthetic route to access 17 and 18.

3.8.2 Reductive amination

The strategy to access the peptide conjugates consisted of a chemoselective ligation step, namely a reductive amination between the two peptide fragments (Scheme 27).



Scheme 27. Reductive amination strategy for the synthesis of the hydrazide mimics.

Preliminary experiments were performed in order to generate the hydrazone-linked peptide. When using aqueous buffers at pH 6 or 7 heating at 50 °C was required, which showed 20% conversion after one day by HPLC. Multiple buffers were tested, as well as aniline catalysis, but none of the strategies showed complete conversion to the product. The fastest ligation conditions found were in water at 50 °C or in MeOH/AcOH 1:1 at room temperature, both of which showed 50% conversion after one day. When it was attempted to purify the hydrazone-linked product, hydrolysis was observed – this could be due to the prolonged exposure to the acidic pH of the buffers. In order to drive the reaction to completion and to increase the stability of the conjugate, it was decided to change the ligation step to a reductive amination and generate a hydrazide linkage instead of the acid-labile hydrazone.

Following on from the ligations performed with the small peptide analogues, the reductive aminations between both peptide fragments were performed in MeOH/AcOH 1:1, in a concentration of 1.5 mM, and with 10 equivalents of $NaBH_3CN$. This afforded **14** in a 36% yield and **(*S*)-15** and **(*R*)-15** in a 43% total yield after purification (the assignment of the diastereomers is presented in Section 3.10.3). The reactions were monitored by HPLC, which in the case of **14** showed complete consumption the aldehyde-containing peptide fragment **17** after only 15 minutes (Figure 101).

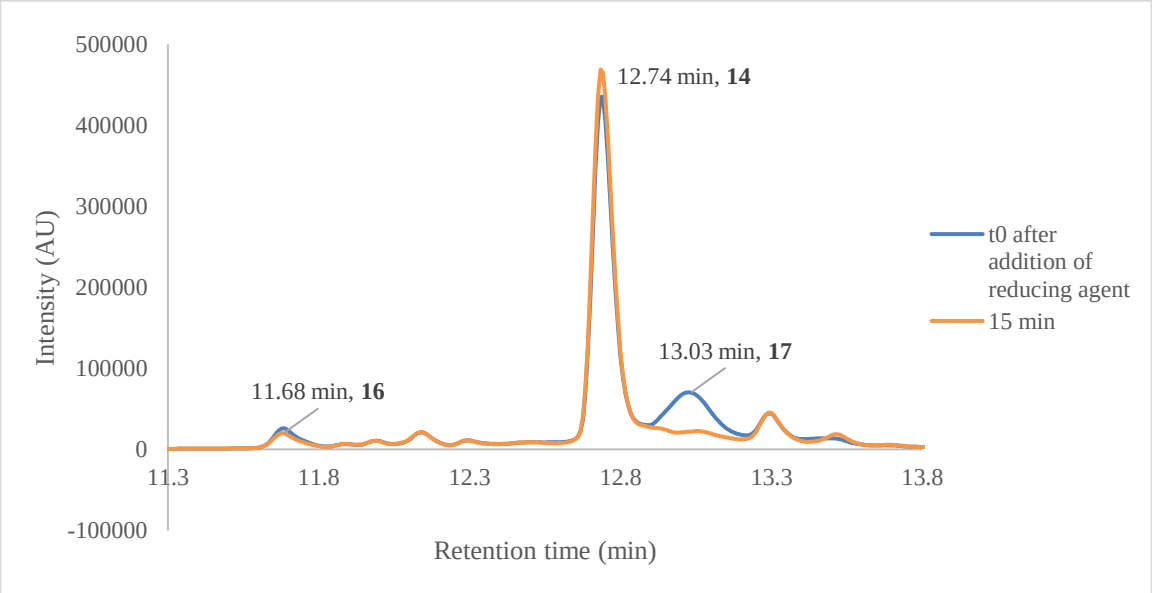


Figure 101. HPLC monitoring at 280 nm of the reductive amination to generate 14.

The HPLC monitoring of the reductive amination towards **(*S*)-15** and **(*R*)-15** (Figure 102) showed that product formation stopped after 10 minutes. In this case, the conversion of the ketone-containing peptide fragment **18** was of 40% after that time.

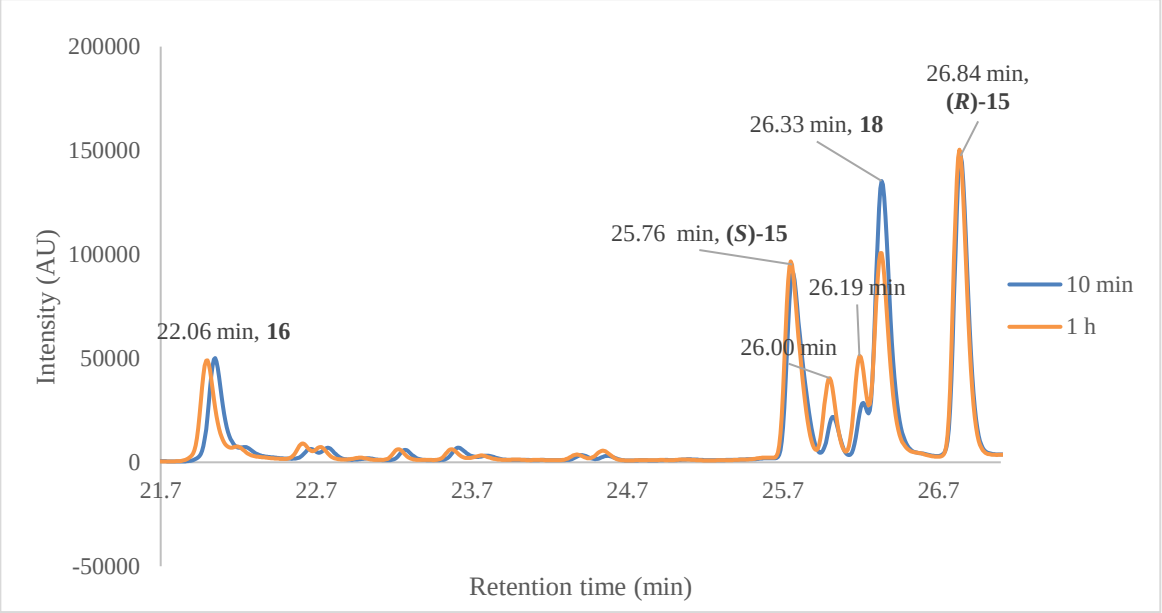
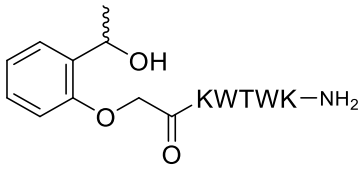


Figure 102. HPLC monitoring at 280 nm of the reductive amination to generate (*S*)-15 and (*R*)-15.

If the reaction was left for longer than 10 min, two by-products were formed instead (peaks at 26.0 and 26.2 min). Subsequent addition of reducing agent further induced the formation of the by-products and not of the desired compounds. Both by-products generated had a m/z of 925.9, which corresponded to the reduction of the ketone in the starting material **18** to generate an alcohol moiety (Scheme 28).



Scheme 28. By-product generated during the reductive amination reaction towards (*S*)-15 and (*R*)-15.

Interestingly, this by-product was not generated during the reductive amination to form **14**. This could be due to the aldehyde starting material **17** reacting faster than **18**, as was observed in the reductive aminations to synthesise the small peptide systems (**19**, **20a** and **20b**). The slow reaction of **18** to form the hydrazide product could compete with its reduction to an alcohol.

The fact that the two diastereomers **(*S*)-15** and **(*R*)-15** were formed in unequal quantities suggests that there could be scope for the development of asymmetric conditions for this ligation.

3.9 Circular dichroism analysis of the peptide conjugates

The peptide conjugates were characterised through circular dichroism (CD) in order to determine their secondary structure and thermal stability. CD is a spectroscopic technique that allows the identification of the folding pattern of a peptide by examining the chirality of its 3-dimensional structure. The far UV absorption spectrum gives information on the secondary structure of the peptide under study, while the near UV spectrum reports on the tertiary structure. The latter further enables the identification of the orientation adopted by aromatic residues, like an edge to face stacking arrangement of Trp residues. In addition, the thermal stability of a construct can be evaluated by following the change in CD as a function of temperature. Although CD provides detailed information on the folding of a peptide, and it can report on the chiral orientation of aromatic residues, it does not report on other specific interactions that take place, like hydrogen bonds. For this reason, NMR analysis has also been employed for the characterisation of these novel systems (see following section).

The CD experiments were run in both the near UV and far UV, with the same conditions that Cochran *et al.*[80] used for the biophysical characterisation of **TrpZip1**: potassium phosphate buffer (20 mM) at pH 7. The concentrations used were 1 mg/mL (0.5 mM) for near and 0.05 mg/mL (22 μM) for far UV.

The far UV CD (Figure 103, MRE was calculated as described in Section 7.12.2) showed that all the peptide conjugates adopted the same conformation as the controls. The negative band at 212 nm was indicative of a β-strand structure, and the positive band at 228 nm was indicative of tryptophan stacking, which confirmed the hairpin conformation.[141,276]

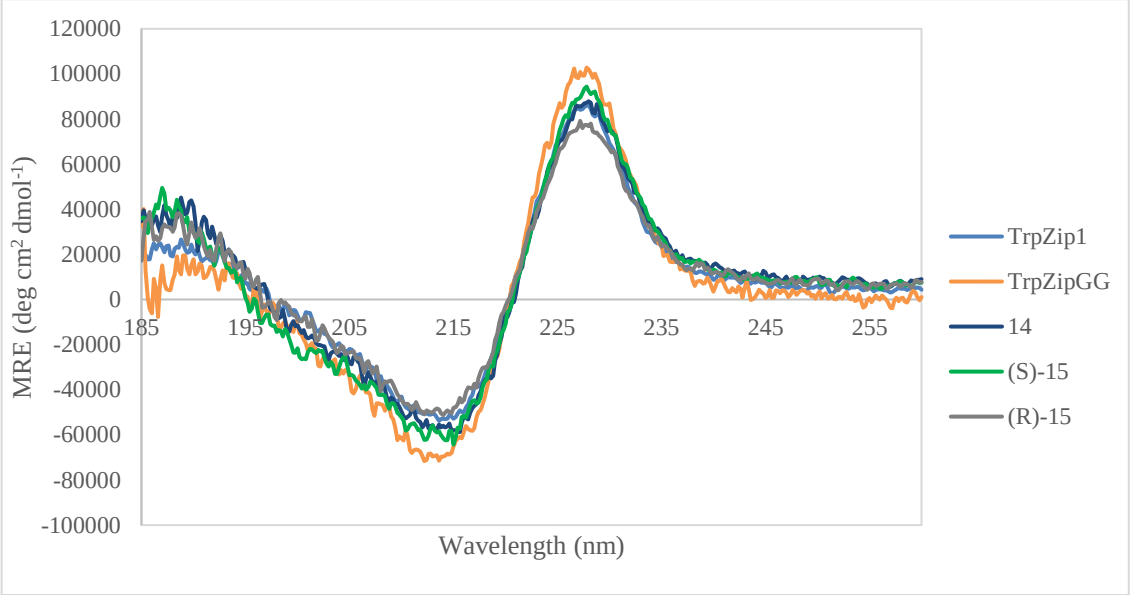


Figure 103. Far UV CD for all the peptide systems in study.

Interestingly, **(*S*)-15** and **(*R*)-15** were not directly superimposable (Figure 104), which showed that the difference in the new chiral centre was causing a slight change in conformation.

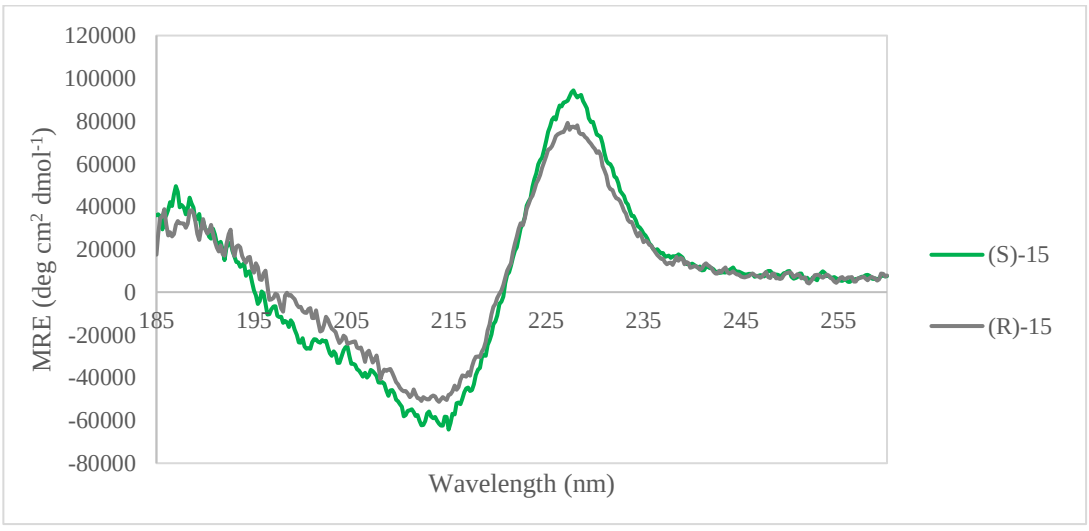


Figure 104. Far UV CD of (*S*)-15 and (*R*)-15.

In the near UV spectra (Figure 105), minima arising from the tryptophan residues were observed at 285 and 295 nm. This confirmed that a defined organisation of the aromatic side chains was taking place. Analogous to what was observed in the far UV CD, both peptides **(*S*)-15** and **(*R*)-15** showed different intensities in their minima, **(*R*)-15** also presenting shallower minima in this case. The same samples were used to run the near and far UV experiments, and their concentrations were measured in the same cell used for CD before and after their thermal denaturation on a UV-Vis spectrometer. Therefore, it was concluded

that the differences between the folding curves of the two diastereomers were not caused by a variation in concentration.

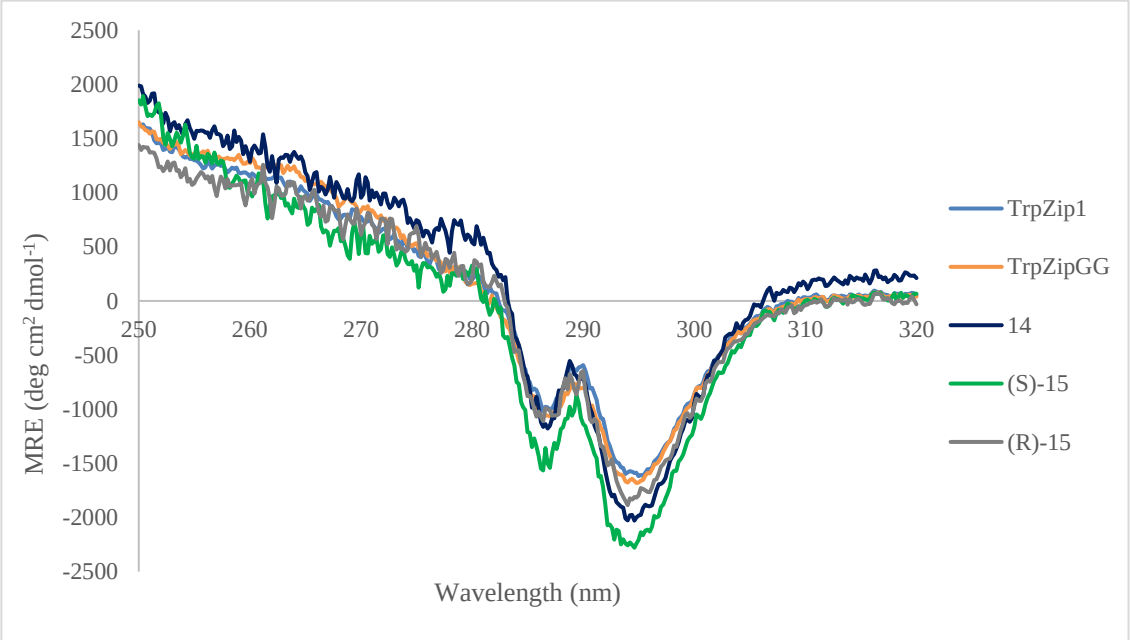


Figure 105. Near UV CD for all the peptide systems in study.

All in all, the CD spectra for the three peptide conjugates was directly comparable to that of the parent **TrpZip1** and **TrpZipGG** peptides, confirming that the β-hairpin structure was maintained.

The thermal stability of the peptide conjugates was assessed by monitoring their change in CD upon gradual heating. The change in mean residue ellipticity (MRE) was monitored at 228 nm as a function of temperature, from 5 °C to 80 °C (Figure 106).

The pre- and post-melt spectra at 5°C were fully superimposable for all peptides, which indicated that the thermal unfolding was completely reversible. The melting curves for the peptide conjugates (Figure 106) were comparable to those of the control peptides, and in comparison to **TrpZip1** the curves were slightly shallower, indicating a slightly higher stability.

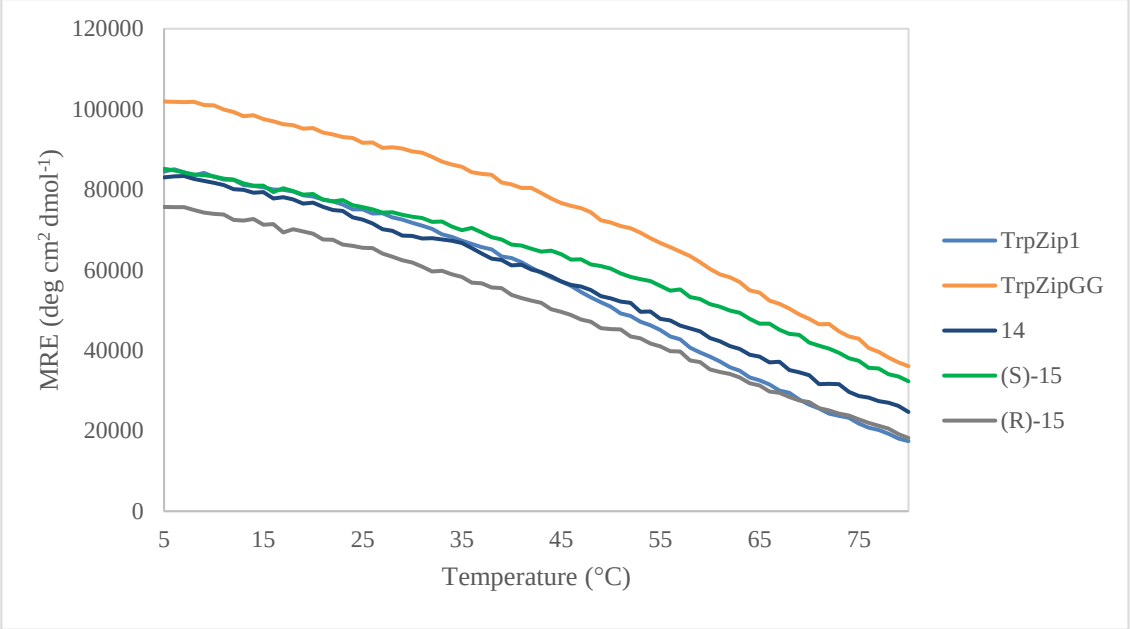


Figure 106. Far UV thermal denaturation experiments at 228 nm for all peptides in study.

Although thermal denaturation curves are typically sigmoidal in proteins, as they fit a two-state folding model, none of the peptides under study showed a clear transition midpoint from the first or second derivatives of the curve (Figure 107).

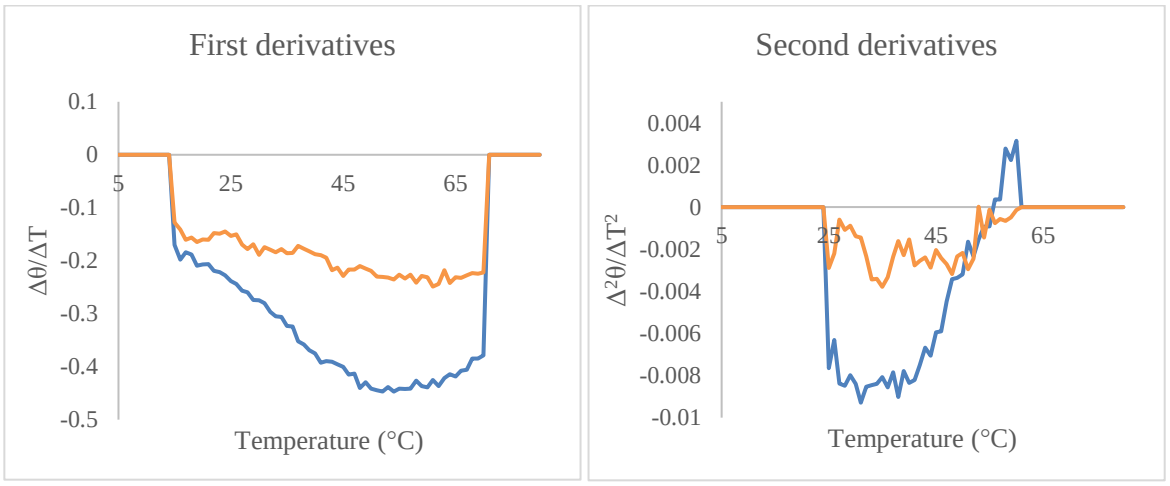


Figure 107. First and second derivatives of the melting curves for TrpZip1 (blue) and 14 (orange).

The melting temperatures (T_m) for these compounds were therefore calculated using the fraction of folding. This was calculated using a two-state unfolding equilibrium (Equation 16):

$$\alpha = \frac{[F]}{[F]+[U]} = \frac{(\theta_t - \theta_U)}{(\theta_F - \theta_U)}$$

Equation 16. Calculation of the fraction folded (α), where [F] and [U] are the concentrations of the folded and unfolded forms, respectively. θ_t is the ellipticity at a specific temperature, θ_U is the ellipticity of the unfolded form and θ_F is the ellipticity of the folded form.[277]

The shapes of the melting curves indicated that the peptides did not reach a completely unfolded state at 80°C, therefore an MRE of zero was adopted as the endpoint for the unfolding process. The MRE values for each peptide at 5 °C were adopted as the fully folded state. The fraction of folding curves obtained for all peptides are presented in Figure 108.

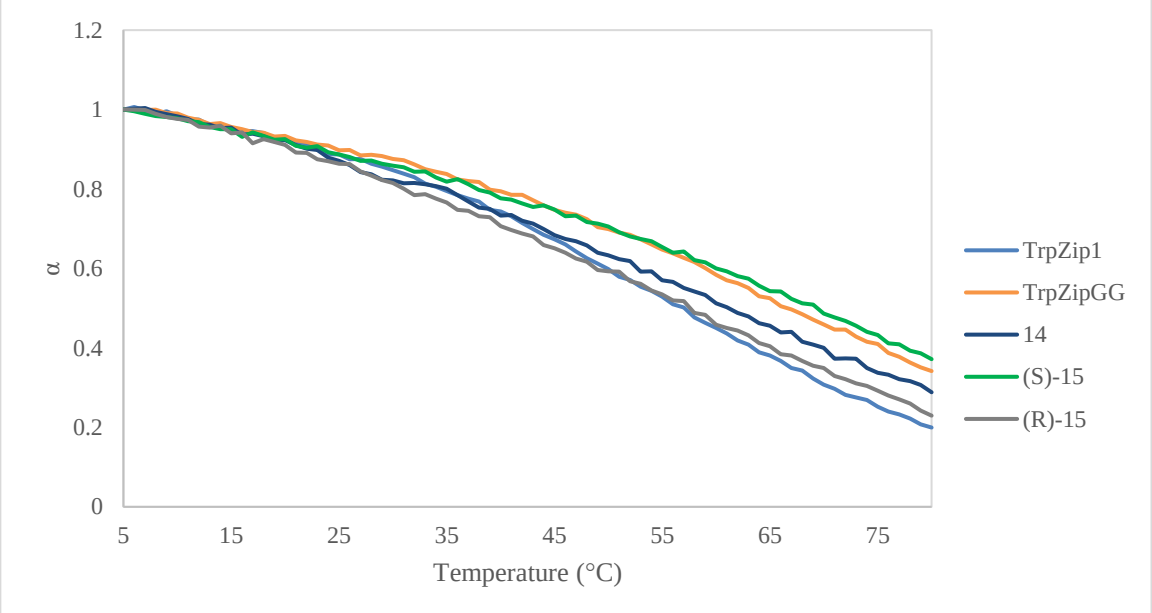


Figure 108. Fraction of folding for all peptides in study.

The calculated T_m were as follows:

Table 12. T_m calculated for the peptides in study.

Sequence	T_m (°C)
TrpZip1	57
TrpZipGG	66
14	62
(*S*)-15	70
(*R*)-15	57

The T_m found for **TrpZip1** was comparable to that calculated by Cochran *et al.*[80] of 50°C. The difference between values could be due to the use of a different method to fit the thermal denaturation data, as the same conditions were used for the thermal denaturation experiments (20 μM peptide concentration in 20 mM potassium phosphate buffer at pH 7). In their work, Cochran *et al.* fit the thermal denaturation curves to a two-state unfolding equilibrium model, allowing the thermodynamic parameters ΔC_p, ΔH_m and ΔS_m to vary. The melting curve of an equimolar mixture of the **Trpzip1** half peptides (SWTWEG and NKWTWK) was recorded and used as the unfolded reference for the fitting.

It was observed that **14** and **(*S*)-15** had an enhanced stability in comparison to the control **TrpZip1**, while **(*R*)-15** maintained the same stability. In addition, **(*S*)-15** also presented a higher stability than the control **TrpZipGG**. Interestingly, the change in conformation of one chiral centre between **(*S*)-15** and **(*R*)-15** caused a considerable difference in stability between the two systems. This could be due to the increased steric hindrance present in the (*R*)-system, which could have a substantial influence in stability in a constrained system like an *ortho*-substituted aromatic ring.

Overall, the β-turn mimics favoured the same folding as the control sequences and the conjugates had a higher stability of the β-hairpin structure compared to the Gly-Asn and Gly-Gly turns of **TrpZip1** and **TrpZipGG**.

3.10 NMR evaluation of the peptide conjugates

Although it is a more time-consuming technique, NMR analysis was utilised in addition to CD because it allows the obtainment of more detailed conformational information. In particular, through-space NOESY experiments enable the confirmation of the hydrogen bonding pattern of the β-hairpin, as well as the disposition of the residue side chains in space. By using this technique, the presence of the *i* to *i+3* like hydrogen bond would be confirmed in the mimic-containing conjugates, and the disposition of the β-turn mimic unit would be studied more closely. In addition, the identity of the **(*S*)-15** and **(*R*)-15** diastereomers would be confirmed.

3.10.1 Control peptides

All resonances were assigned for the two control peptides using 2D COSY, TOCSY, NOESY and HSQC spectra. The experiments were run using the conditions from Cochran *et al.*,[80] at 2 mM and 1 mM peptide concentrations, in an AcOH-d_4 buffer (10 mM, pH 5.5) with 5% D_2O and using DSS as an internal standard, at 288 K. The NMR experiments of the control sequences and their preliminary assignment were performed in collaboration with Dr Albane Neckebroeck.

The residue connectivity was established using the sequential NOEs between H^{α}(Xaa)–NH(Xaa+1), which are very strong in antiparallel β-sheets.[278,279] The β-hairpin conformation was confirmed *via* cross-strand NOEs between H^{α} protons in non-hydrogen bonded positions and between NH protons in hydrogen bonded positions (Figure 109).

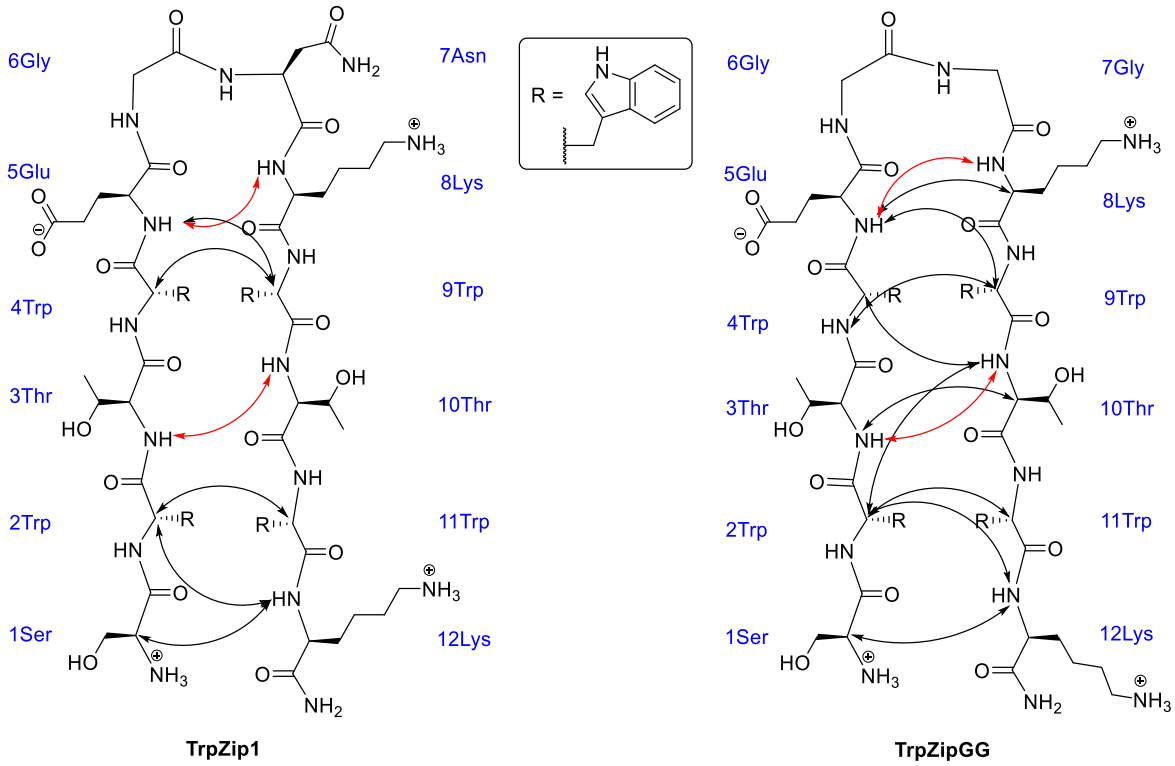


Figure 109. Cross-chain NOEs found for TrpZip1 (left) and TrpZipGG (right). The NH to NH cross peaks are represented in red.

The turn conformation was confirmed for both peptides (Figure 110): most importantly, the *i* to *i+3* hydrogen bond was confirmed with an NOE between 5Glu NH and 8Lys NH. In addition, cross peaks were observed between the NH and H^{α} protons within the turn, demonstrating their close proximity.

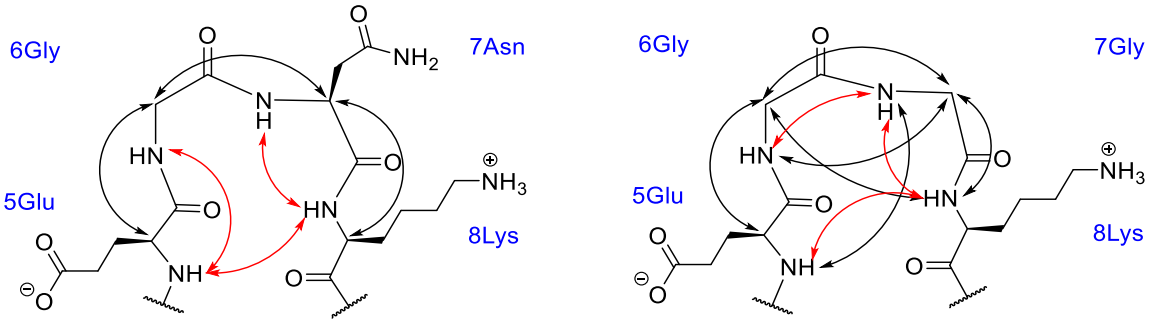


Figure 110. Selection of in-turn NOEs observed for TrpZip1 (left) and TrpZipGG (right). The key cross peaks for determining the conformation are depicted in red.

The edge to face stacking between Trp residues was observed for both peptides. NOEs were present between the aromatic protons in 4Trp and the H^{α} in residues 6 and 7, confirming that said tryptophan was behind the β-turn. An extreme upfield shift was observed for the $H^{\varepsilon 3}$ protons in 4Trp and 11Trp (5.5 and 5.4 ppm for **TrpZip1** and 5.7 and 5.3 ppm for **TrpZipGG**), which was caused by their contribution to the edge to face stacking.[280] This effect is explained by the ring current field that takes place in aromatic rings (Figure 111), which is generated by the electrons that circulate in the π system. This causes the protons that are within that ring current to be more shielded and experience an upfield shift.[281]

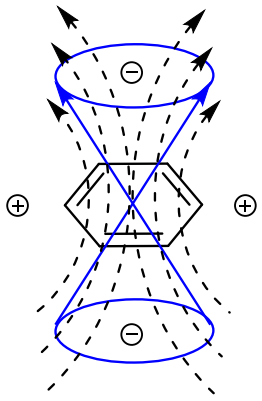


Figure 111. Ring current field generated by benzene. The blue double-cone represents the shape of that field, while the dashed lines are the magnetic field. The minus signs indicate that the atoms located inside of the cone in the protein structure are shifted upfield, whereas the atoms outside the cone are shifted downfield (plus signs).

As it can be observed in the 3D NMR structure of **TrpZip1** calculated by Cochran *et al.*[80] (Figure 112), the tryptophan residues stack in an edge to face manner, where the $H^{\varepsilon 3}$ protons in 4Trp and 11Trp act as the 'edge' component in the interaction. They point towards the centre of the subsequent tryptophan indole ring, therefore experiencing its ring current, which causes the upfield shift of their $H^{\varepsilon 3}$ protons.

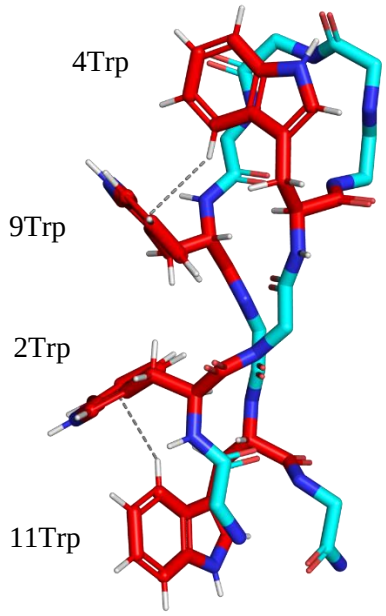


Figure 112. 3D NMR structure of TrpZip1 (PDB 1LE0), where the Trp residues are represented in red. From the hairpin termini to the turn, the stacking order is 11Trp/2Trp/9Trp/4Trp. The interactions between the $H^{\varepsilon 3}$ protons in 4Trp and 11Trp and the subsequent Trp indole are represented in grey.

Because the chemical shifts for the 4Trp $H^{\varepsilon 3}$ and 11Trp $H^{\varepsilon 3}$ protons in **TrpZipGG** were comparable to those in **TrpZip1** and the same cross peaks between side chains were observed, it could be inferred that the same edge to face stacking pattern of the tryptophan residues was taking place. All in all, both control peptides adopted the same β-hairpin structure, with the same turn and side chain conformations.

3.10.2 NMR analysis of 14

2D NOESY, COSY and TOCSY spectra were used to assign all resonances in **14**, at both 278 and 288 K. Experiments were run at a 1 mM concentration, in AcOH-d_4 buffer (10 mM, pH 5.5) with 5% D_2O and using DSS as the internal standard. The connectivity of the U-BTM unit to the rest of the backbone was confirmed (Figure 113) by the presence of NOE cross peaks between both protons in {1} and {2}(a, b) as well as between both protons in {3} and{2}(c, d). In this nomenclature the spin systems are indicated in braces, and the resonances within the spin system are indicated in parentheses. Both protons in the spin systems {1} and {3} were diastereotopic and related by a COSY cross peak. The protons in {3} also correlated to 7Lys NH through an NOE cross peak. Interestingly, an NOE correlation was observed between both protons in {3} and one of the protons in {1}, indicating a compact structure. This matched what was observed in the MD simulations

(Figure 113), where one of the protons in {1} was closer to the spin system {3} than the other.

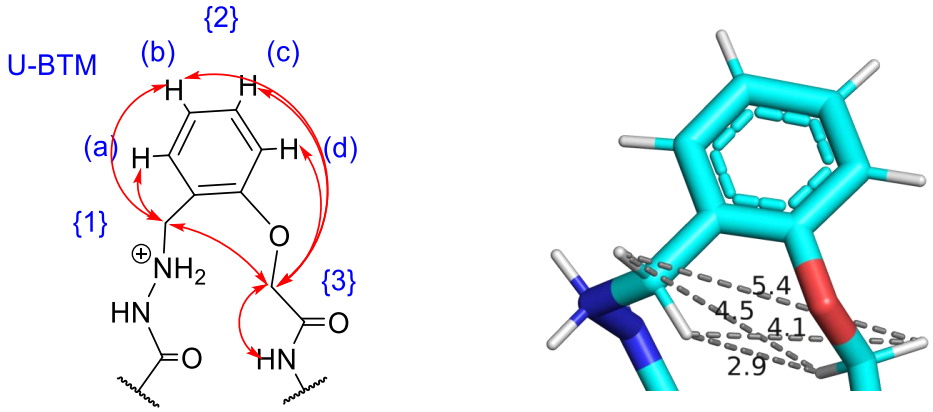


Figure 113. Left: NOE correlations observed within the U-BTM unit in 14. Right: frame from the MD simulation showing the distances between the protons in {1} and {3}.

The differentiation and assignment of 7Lys and 11Lys (Figure 114) was complex, as their chemical shifts were very similar. Experiments at 278, 283 and 288 K were run in order to confirm the presence of the 7Lys NH proton (Figure 115). It was observed that as the temperature increased, the distortion of the doublet (11Lys NH) decreased and a new, broadened signal appeared, corresponding to 7Lys NH. In addition to confirming its presence, these experiments also highlighted how the amide proton in 7Lys was exchanging with solvent more than that of 11Lys, which indicated that the structure was dynamic and a weaker hydrogen bond was taking place with this residue.

Another difference between the lysine residues lay in their side chains – the β, γ, δ and ε protons in 11Lys were diastereotopic, indicating a constrained conformation. On the contrary, the side chain protons in 7Lys were all degenerate, suggesting a free moving side chain.

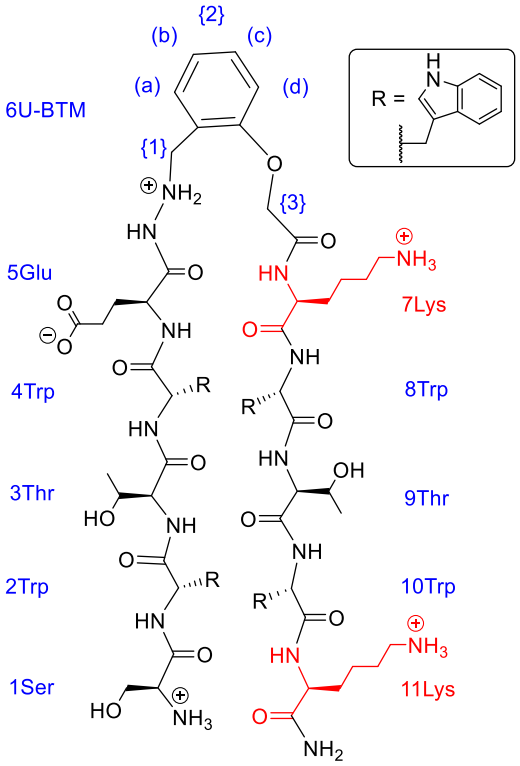


Figure 114. Structure of 14, showing 7Lys and 11Lys in red.

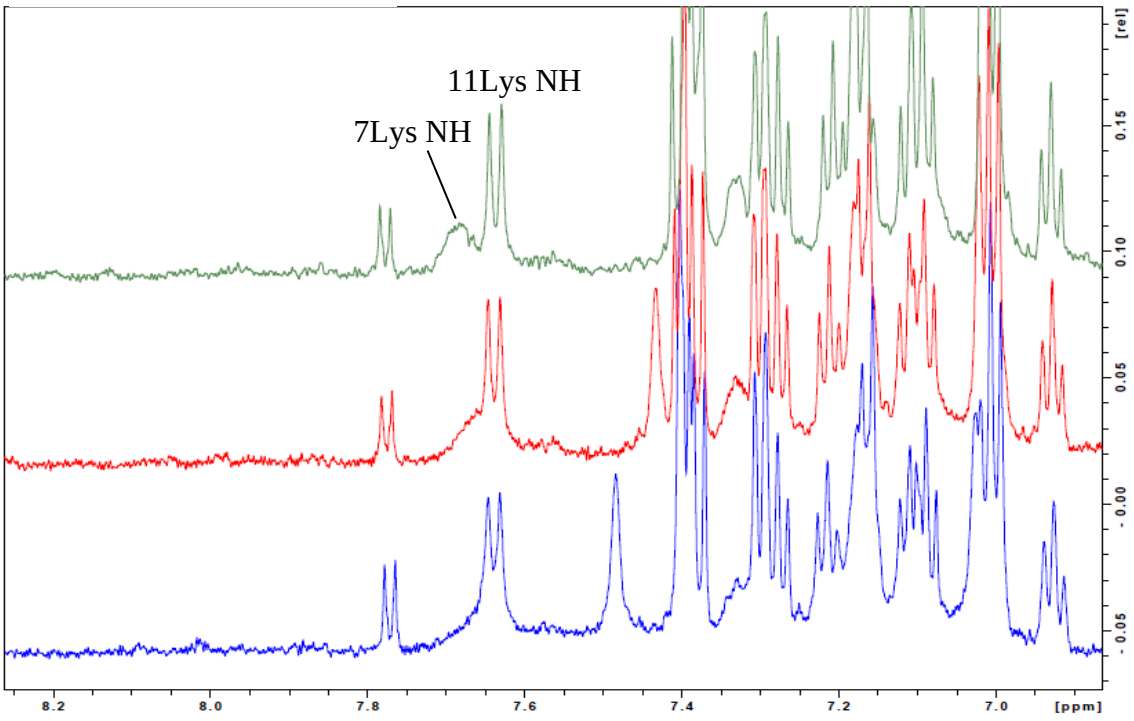


Figure 115. Temperature-dependent 1D ^{1}H NMR experiments for 14. From bottom to top, the temperatures are 278, 283 and 288 K. ^{1}H NMR experiments performed at a 1 mM peptide concentration, in AcOH-*d*$_4$ buffer (10 mM, pH 5.5), in 5% D_2O on a 600 MHz spectrometer.

The β-hairpin structure was confirmed by the presence of intense H^{α}(Xaa)–NH(Xaa+1) cross peaks as well as the chemical shifts of the NH and H^{α} protons, which were commensurate with a β-strand secondary structure.[278,279] NOE cross peaks were observed between amide NH groups (Figure 116), highlighting the cross-sheet hydrogen bonded structure. Most importantly, the through-space interaction between the 5Glu NH to 7Lys NH confirmed the presence of the β-turn-like hydrogen bond. Experiments in 100% D_2O were used to identify the H^{α} to H^{α} NOEs, which showed that those interactions were evident for the non-hydrogen bonded positions.

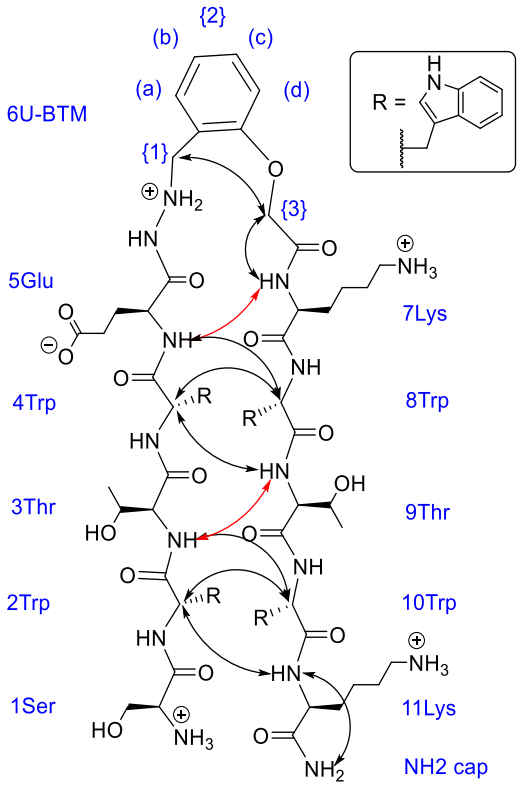


Figure 116. Across-chain NOEs found for 14. The NH to NH cross peaks are represented in red.

The dispositions of the side chains were conserved between **14** and the control sequences, and the same ordering of side chains was observed within the aromatic stack. The chemical shifts of the $H^{\varepsilon 3}$ protons in 4Trp and 10Trp were considerably shifted upfield (5.9 and 5.4 ppm, respectively), confirming their participation in an edge-to-face stacking interaction and consistent with what was observed in the control peptides.

3.10.3 NMR analysis of (*S*)-15 and (*R*)-15

Initial experiments were run at 288 K for **(*S*)-15**, consistent with the other sequences under study, but the H^{α} protons as well as the 7Lys NH could not be identified at this temperature and therefore the main chain connectivity could not be confirmed. Therefore, 1D ^{1}H NMR experiments were run at different temperatures (Figure 117). These experiments showed broadening of the NH and H^{α} signals at lower temperatures, which could indicate the flipping of side chains between different orientations. The resolution of the NH and the H^{α} region was improved at 298 K, and therefore the 2D experiments were run at this temperature, at a 0.7 mM concentration (in AcOH-d_4 buffer 10 mM, pH 5.5, 5% D_2O/DSS). Experiments in 100% D_2O were also used at this temperature to reveal the H^{α} to H^{α} NOE cross peaks.

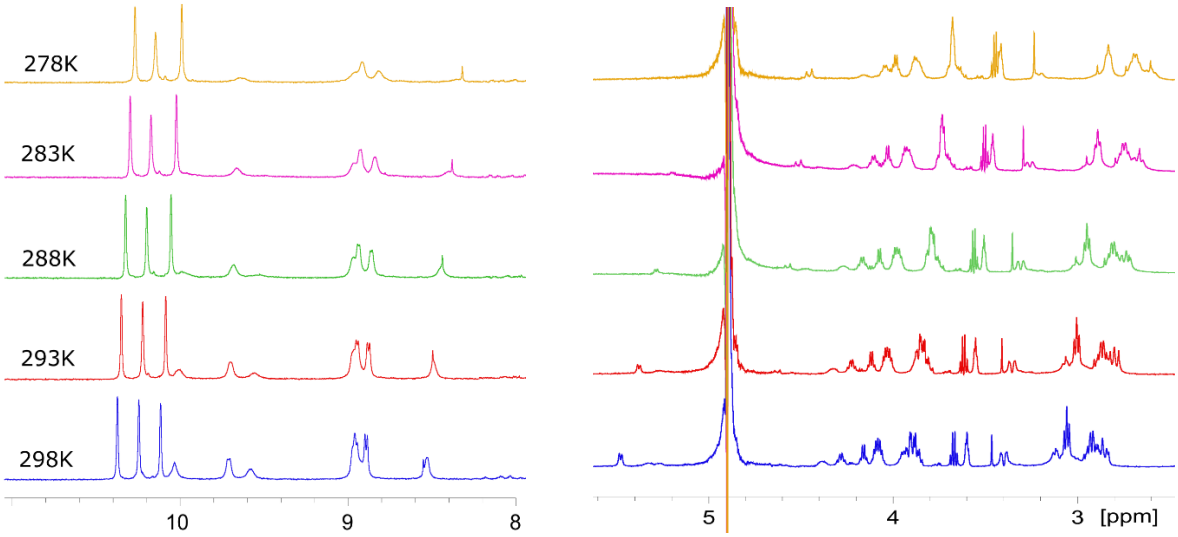


Figure 117. Temperature-dependent 1D ^{1}H NMR experiments for (*S*)-15. The NH region is shown on the left and the H^{α} region on the right. ^{1}H NMR experiments performed at a 0.7 mM peptide concentration, in AcOH-d_4 buffer (10 mM, pH 5.5), in 5% D_2O on a 600 MHz spectrometer.

In the case of **(*R*)-15**, NMR experiments were run at a concentration of 0.6 mM in AcOH-d_4 buffer (10 mM, pH 5.5), 5% D_2O/DSS and at 288 K.

The connectivity of the turn unit to the rest of the chain in both peptides was confirmed by NOE cross peaks between the aromatic system {2} and the diastereotopic CH_2 {3}, as well as between {2} and the CH_3 ME and CH HA in spin system {1} (Figure 118).

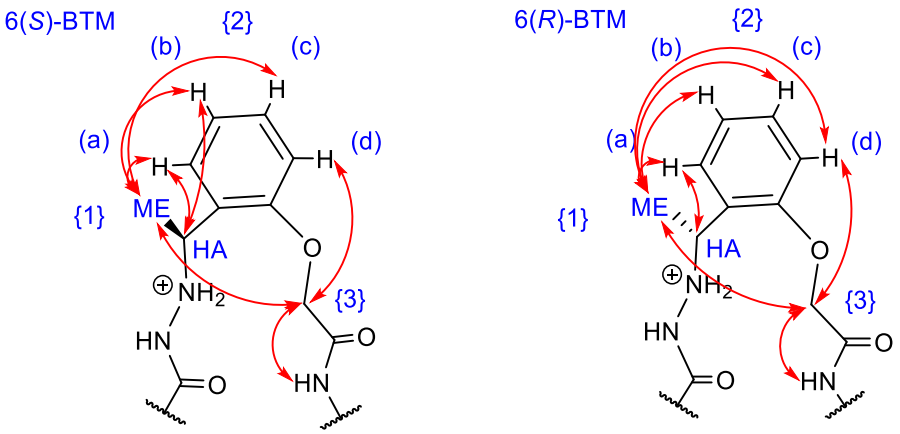


Figure 118. In-turn NOEs for (*S*)-15 (left) and (*R*)-15 (right).

Regarding the β-hairpin structures, cross-sheet NH to NH cross peaks were observed, which confirmed the presence of the β-turn-like hydrogen bond (Figure 119). H^{α} to H^{α} NOEs were also observed between non-hydrogen bonding positions – in the case of **(*R*)-15**, the 4Trp H^{α} to 8Trp H^{α} NOE was not observed due to those signals being under the solvent peak. In addition, intense H^{α}(Xaa)–NH(Xaa+1) cross peaks took place and the chemical shifts of the NH and H^{α} protons corresponded to a β-strand secondary structure.

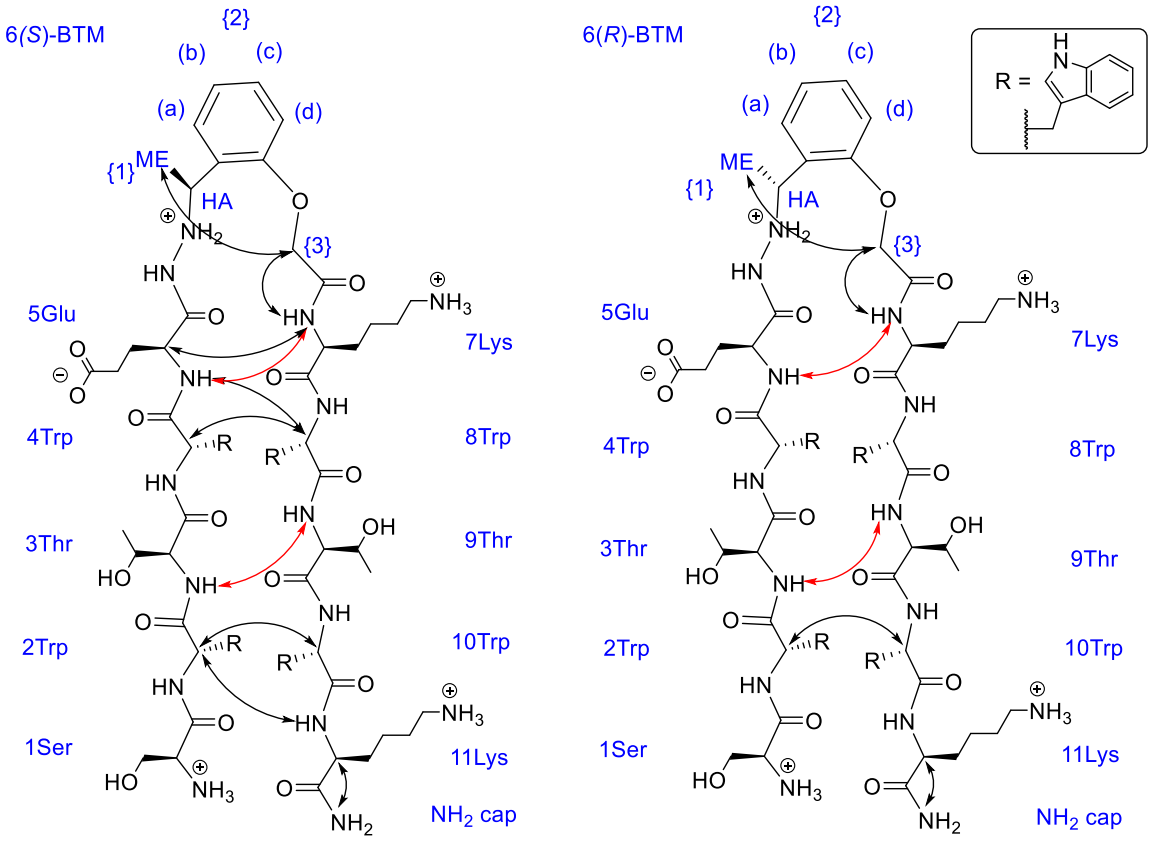


Figure 119. Cross-chain NOE cross peaks for (*S*)-15 (left) and (*R*)-15 (right). NH to NH cross peaks are represented in red.

The same side chain disposition in space was observed for these conjugates as in the control sequences. The edge-to-face stacking between the Trp residues was confirmed by the NOE pattern as well as the upfield shift of the $H^{\varepsilon 3}$ protons in 4Trp and 10Trp (5.90 and 5.48 ppm for **(*S*)-15** and 6.45 and 5.39 ppm for **(*R*)-15**, respectively).

In order to identify which diastereomer corresponded to **(*S*)-15** and **(*R*)-15**, the distances between spin system {1} and neighbouring groups were used. To this end, the presence or absence of an NOE was compared to the average distance extracted from the MD simulations (Figure 120). The distances between {1}ME and {3} and between {1}ME, HA and 4Trp were chosen for comparison (NOE cross peaks between {1}HA and {3} were not observed for either system).

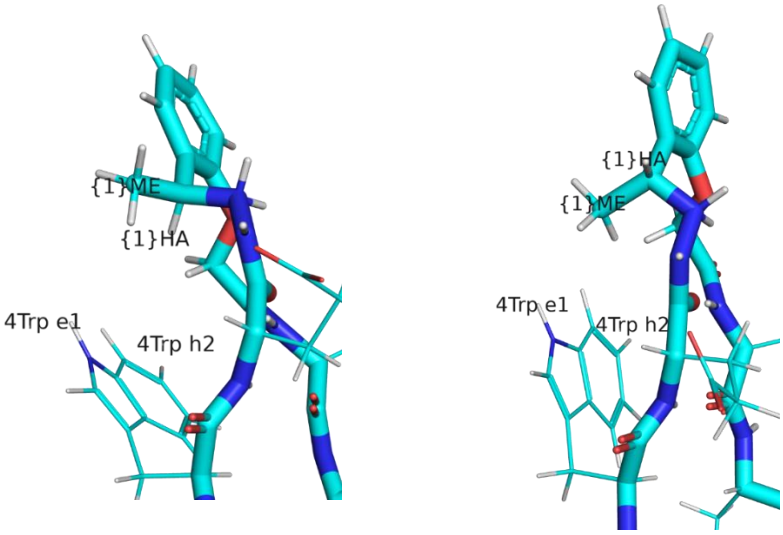


Figure 120. Representative frames from the MD simulations of (*S*)-15 (left) and (*R*)-15 (right), showing the close proximity of {1}ME,HA to 4Trp.

The average distances extracted from the MD simulations and the presence or absence of NOE cross peaks for each contact are presented in Table 13. It was observed that although the MD simulations predicted a significant difference in distance between {1}ME and spin system {3}, an NOE cross peak was present in both peptides. The key contacts for assignment were between $4Trp^{\eta 2}$ and $4Trp^{\varepsilon 1}$ to {1}HA, as an NOE cross peak was observed between the two pairs for **(*S*)-15**, but not for **(*R*)-15**. This constituted strong evidence that **(*S*)-15** corresponded to the (*S*)-diastereomer.

Table 13. Average distances from MD simulations and presence or absence of NOE cross peaks for each contact.

Average distances extracted from MD trajectories (Å)		**NOE cross peaks**	
(*S*)-15	**(*R*)-15**	**(*S*)-15**	**(*R*)-15**
{1}ME (carbon) – {3}*pro*-R 3.3, {3}*pro*-S 4.4	{1}ME (carbon) – {3}*pro*-R 5.3, {3}*pro*-S 5.1	Observed	Observed
{1}ME (carbon) – 4Trp$^{\varepsilon 1}$ 3.5	{1}ME (carbon) – 4Trp$^{\varepsilon 1}$ 3.7	Observed	Observed
4Trp$^{\eta 2}$ – {1}HA 6.7	4Trp$^{\eta 2}$ – {1}HA 7.9	Observed	Not observed
4Trp$^{\varepsilon 1}$ – {1}HA 3.4	4Trp$^{\varepsilon 1}$ – {1}HA 3.8	Observed	Not observed

Further confirmation of the assignment was sought from the peak heights for the NOE cross peaks observed (Table 14). The peak heights corresponding to the NOE between {1}ME and the two protons in {3} were consistent with the MD distances predicted for **(*S*)-15**, as one of the cross peaks had a higher intensity than the other. Regarding **(*R*)-15**, the two peak heights between {1}ME and {3} had similar values to each other, which was in accordance with the MD distances, but the peak intensities were similar to those of **(*S*)-15**. This contradicted what was expected, as the longer distance between {1}ME and {3} in **(*R*)-15** would have caused the peak height to be lower.

More information was obtained by analysing the peak heights between {1}ME and 4Trp$^{\varepsilon 1}$. Although both diastereomers showed an NOE cross peak between them, their intensities were significantly different: the MD simulations predicted a difference in the average distances of 0.2 Å, but the peak heights differed by an order of magnitude. This further supported that **(*S*)-15** corresponded to the (*S*)-diastereomer and indicated that the MD simulations underestimated that distance.

Table 14. Peak heights of the NOE cross peaks observed for (*S*)-15 and (*R*)-15.

	(*S*)-15	**(*R*)-15**
{1}ME – {3}	33325.5, 40796.8	35534.5, 35675.6
{1}ME – 4Trp$^{\varepsilon 1}$	177445.5	22720.3
4Trp$^{\eta 2}$ – {1}HA	30209.8	Not observed
4Trp$^{\varepsilon 1}$ – {1}HA	248162.5	Not observed

3.11 Conclusions and future work

Three β-turn mimics have been rationally designed and evaluated as Ala-flanked tripeptides as well as within a β-hairpin backbone. The tripeptides were used to explore suitable reductive amination conditions, and to compare the β-hairpin conjugates to a system that would be unfolded without the presence of a turn-inducing unit.

A ligation protocol has been established to install the β-turn mimic unit at the same time as ligating together two peptide fragments under mild conditions, which allows the use of unprotected peptide substrates. CD analyses have shown that the hairpin structure is maintained in the peptide conjugates, and that the thermal stability of the conjugates is slightly higher than that of the original **TrpZip1** control. NMR evaluation of the mimics further confirmed that the conformation adopted by the conjugates was the same as that of the control sequences, indicating that the key hydrogen bond and subsequent geometry of a β-turn were present. Overall, the β-turn mimics operate as designed, resulting in a hydrogen bonded β-turn-like structure that does not perturb either the backbone or side chain conformations.

Future work could involve the development of asymmetric conditions for the reductive amination step, as the diastereomers **(*S*)-15** and **(*R*)-15** are formed in unequal quantities and they also have different thermal stabilities. Another direction would be to adapt the ligation protocol to perform the reductive amination in aqueous buffers, in order to widen its applicability to proteins that would need to be kept in folding conditions.

Another possibility would be the use of this ligation protocol in larger systems, in order to evaluate what effects the introduction of a mimic unit could have in a system where contacts between residues that are far apart in the sequence take place.

4 Design of an MDM2/p53 interaction inhibitor

The work presented in this chapter was performed in collaboration with Prof Danny Huang's group at the Beatson Institute. Dr Mads Gabrielsen performed the SPR assays.

4.1 Protein-protein interactions

Protein-protein interactions (PPIs) regulate nearly all cellular functions, with thousands of them having been characterised in humans. They mediate many regulatory pathways, and are therefore involved in the development of many diseases.[282] Due to their role in the development of disease, and in particular of cancer, these interactions have attracted substantial attention as targets for inhibitor development. PPIs are difficult to target because they take place over large surface areas (of 1000 to 4000 $Å^2$, Figure 121) and because they are dynamic due to conformational changes that take place upon association. [283,284]

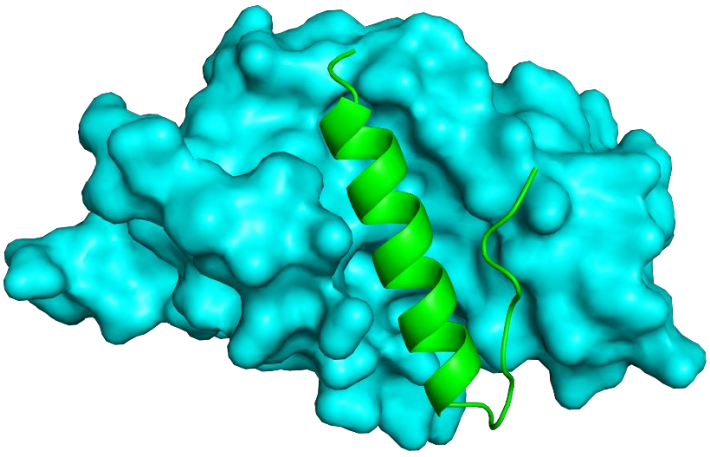


Figure 121. PPI between the Bcl-xL protein and the BIM peptide. PDB code 1PQ1.

PPIs take place through 'hot spot' residues which contribute the majority of the binding affinity and specificity to the interaction. Due to the contact area being large, small molecules usually have low affinity for these interactions and medium-sized inhibitors are better suited scaffolds for competitive inhibition.[285] This has focused the attention on peptides, which offer advantages such as the direct similarity to protein fragments, ease of synthesis and lower toxicity upon degradation in comparison to small molecule inhibitors, as they degrade into amino acids.[286] Although these are attractive targets, peptides have certain disadvantages, mainly their low proteolytic stability due to the susceptibility to hydrolysis of the peptide bonds caused by proteases,[287] and also their low conformational stability in short, linear peptides. In order to overcome these drawbacks, non-natural

modifications can be used. These include the modification of the peptide backbone through the use of peptoids, β-amino acids and D-amino acids, as well as cyclisation techniques, which include the use of hydrogen bond surrogates, stapling strategies and cyclic β-hairpins.[282]

A significant number of PPIs involve α-helices,[93] and many inhibitors have been designed that adopt this secondary structure[288] or mimic it.[94] One particular strategy for the mimicry of these structures is the use of a β-hairpin template, which displays the key interacting residues in the correct orientation for binding. Most of the hairpin backbones used are cyclic and contain turn-inducing units, which increase their conformational stability.[282] The most commonly used β-turn unit is the (D-Pro)-(L-Pro) motif,[160,179,180] although ornithine[181] and other unnatural units like benzodiazepine[182] (Figure 122) and *trans*-pyrrolidine-3,4-dicarboxamide[289] scaffolds have also been used.

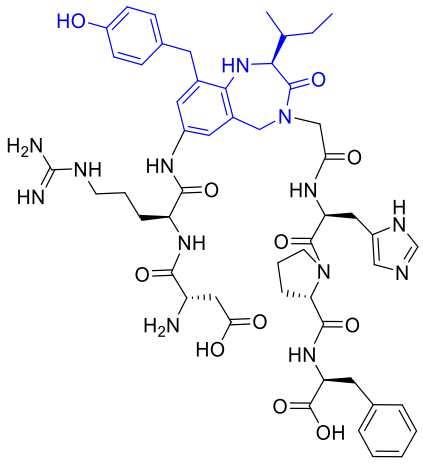


Figure 122. β-Hairpin containing a benzodiazepine-based β-turn mimic (in blue) which inhibits the angiotensin II receptors.[182]

4.2 The MDM2/p53 interaction

p53 is a tumour suppressor that prevents cells from malignant transformation by inducing cell cycle arrest or apoptosis when DNA damage or cellular stress occurs. The murine double minute 2 (MDM2) and its human homologue HDM2 protein are ubiquitin E3 ligases that regulate p53 levels by means of a PPI that neutralises the activity of p53. In healthy cells the p53 levels are kept low and are downregulated by MDM2, but in tumours loss of activity of the p53 protein can happen due to the overexpression of MDM2. This is a major contributor to cancer development,[290] and therefore significant efforts have been made to develop inhibitors of the MDM2/p53 interaction to recover p53 activity.

The MDM2/p53 interaction takes place between the α-helical domain of p53 and a hydrophobic pocket on the MDM2 surface (Figure 123).[291] There are three hot spot residues in p53 that are essential for binding, namely 19Phe, 23Trp and 26Leu, which insert into the MDM2 cleft.[292]

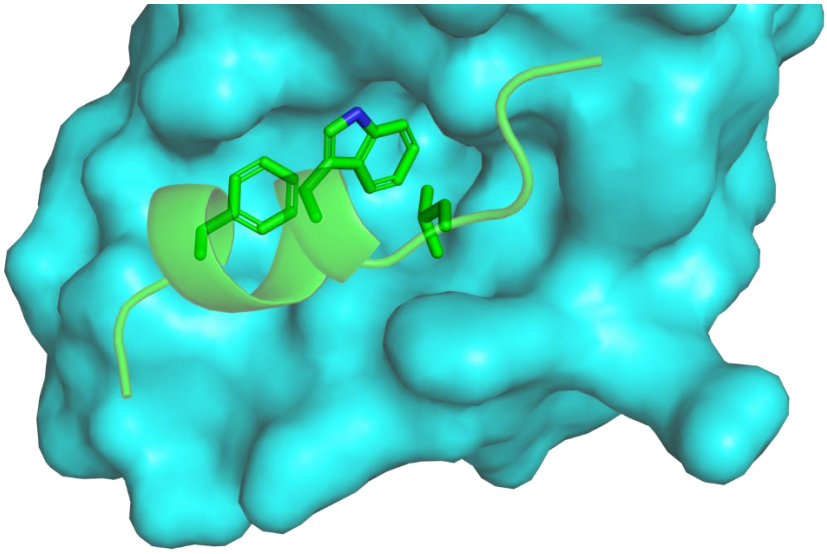


Figure 123. Crystal structure of the HDM2 protein (in blue) bound to the α-helical domain of p53 (in green). The hot spot residues in p53 are represented as sticks. PDB code 1YCR.

Many inhibitors of this interaction have been previously developed, which include peptoids,[293] α-helices stabilised with hydrogen bond surrogates,[294] mini proteins[295] and stapled α-helices.[292,296,297] The latter is the most explored approach, as it highly stabilises the α-helical structure of the inhibitor and therefore helps overcome the usually low proteolytic resistance of small peptide therapeutics. For example, Baek *et al.*[298] have designed an α-helix which contains an all-hydrocarbon staple between two non-natural residues that are seven residues apart (Figure 124). They found that the peptide oriented the three hot spot residues in the correct geometry, and in addition the staple interacted with the protein, which the authors suggested enhanced the binding affinity. Interestingly, it was observed that one of the hot spot residues, 26Leu, adopted a slightly different conformation, which might suggest that the structure was too constrained.

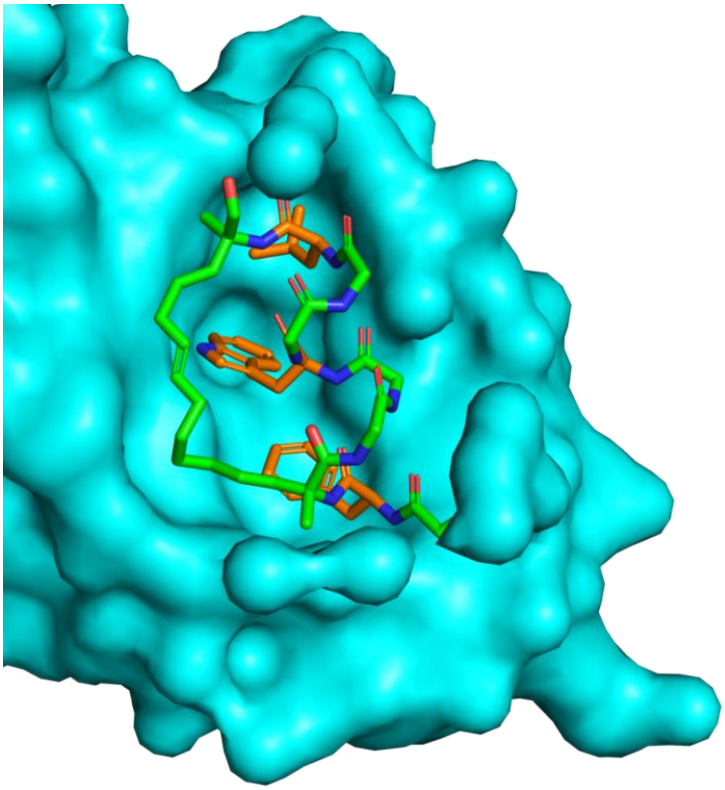


Figure 124. Crystal structure of the stapled inhibitor designed by Baek *et al*.[298] (in green) bound to MDM2 (in blue). The hot spot residues are represented in orange. PDB code 3V3B.

4.3 Inhibition of the MDM2/p53 interaction using a β-hairpin

One of the few examples of PPI inhibitors with a β-sheet structure was designed by Fasan *et al*.,[160] who used the protein epitope mimetic approach[127] to graft the hot spot residues in p53 onto a β-hairpin backbone (Figure 125). The authors hypothesised that the hot spot residues could be oriented in the correct conformation by a hairpin scaffold, as the distances between the C^{α} atoms of 19Phe and 23Trp and that of two *i* and *i+2* residues along a β-hairpin strand were very similar.

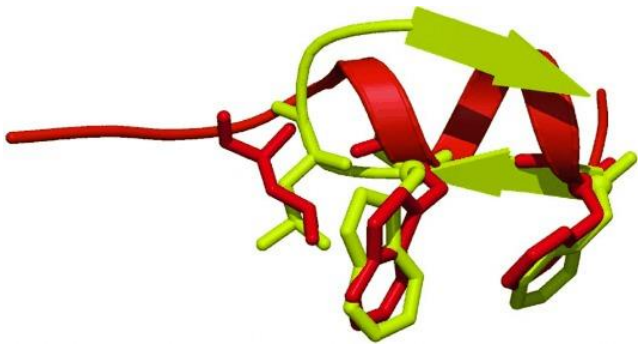
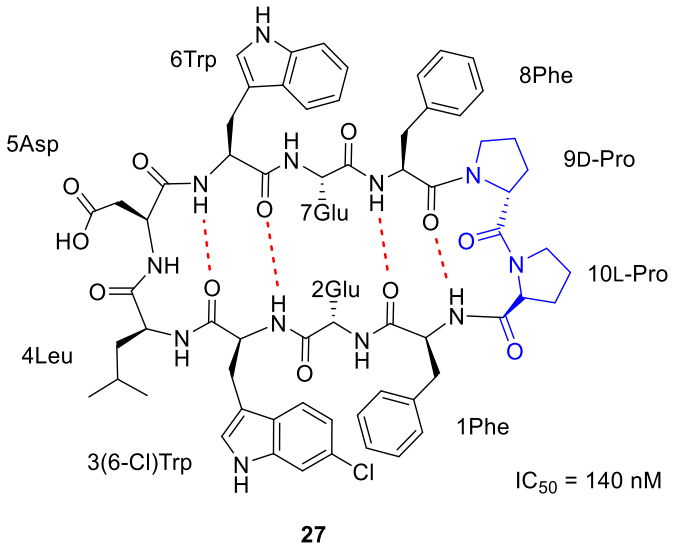


Figure 125. Superimposed α-helical domain of p53 (red) and a model β-hairpin (yellow). Reproduced with permission from ref.[160]

A cyclic β-hairpin was used as the scaffold, with a (D-Pro)-(L-Pro) turn in order to stabilise the conformation further. A library of peptides was synthesised, where various amino acids along the sequence were altered, including an alanine scan. The best hit found using natural amino acids along the β-hairpin strand had an IC_{50} of 0.53 ± 0.06 μM, and when the Trp residue was substituted by (6-Cl)-Trp (**27**, Figure 126) the IC_{50} improved to 0.14 ± 0.06 μM.



27

Figure 126. Lead mimetic identified,[299] containing a (6-Cl)-Trp residue and the (D-Pro)-(L-Pro) turn (in blue). The hydrogen bonds are represented in red.

A subsequent study by the same group focused on an SAR, which confirmed that the previously identified (6-Cl)-Trp-containing hairpin **27** was the most active inhibitor.[299] A crystal structure of the inhibitor bound to HDM2 was obtained (Figure 127, top), which confirmed that binding was taking place through 1Phe, 3(6-Cl)-Trp and 4Leu. It was also observed that the side chain of (6-Cl)-Trp pointed deep into the hydrophobic HDM2 binding pocket (Figure 127, bottom) and established a non-covalent interaction with 86Phe in HDM2, which caused the increase in binding affinity. On the other β-hairpin strand, an edge to face interaction was observed between 6Trp, 8Phe and 55Phe in HDM2, while the 2Glu and 7Glu side chains were in solvent-exposed positions. Interestingly, the amino acids in the unnatural turn unit were found to not participate in binding.

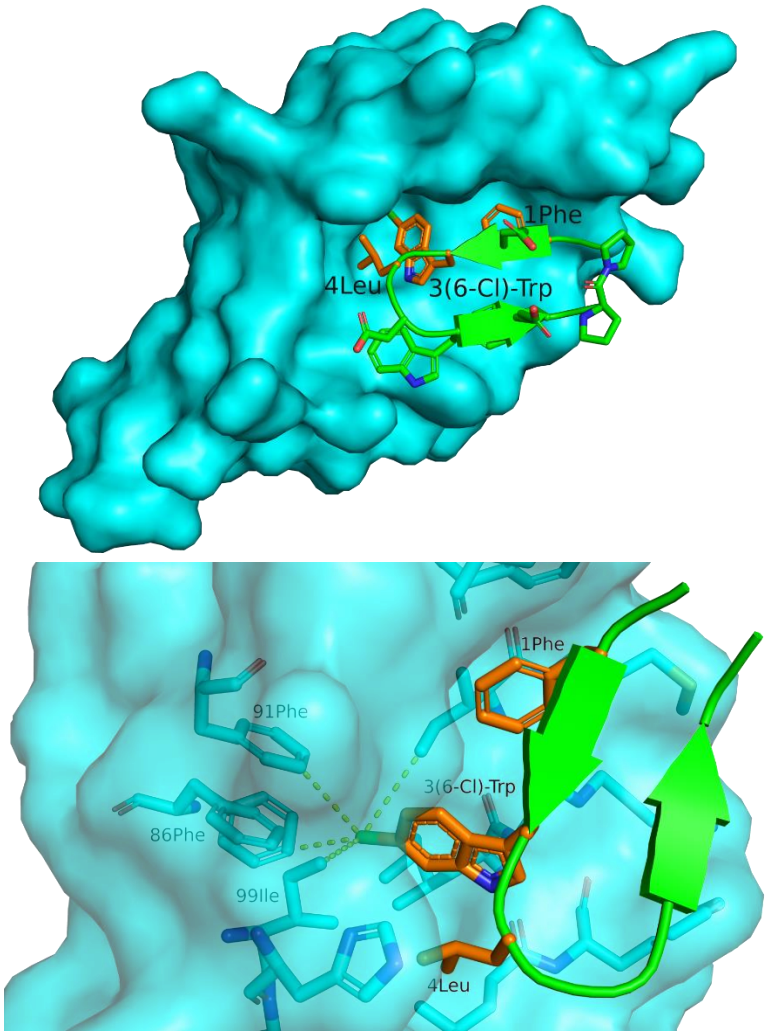


Figure 127. Top: crystal structure of inhibitor 27 (in green and hot spot residues in orange) bound to HDM2 (in blue). Bottom: Close-up of the interactions established by 3(6-Cl)-Trp. PDB code 2AXI.

Fasan *et al*. identified the higher stability of the β-hairpin structure as one of the reasons for the increased affinity of this inhibitor in comparison to other mutants in the library. But questions have since been raised on whether more rigidity will always equate to a higher affinity. Danelius *et al*.[300] performed NMR analyses on this inhibitor and compared it to three β-hairpin mutants containing a (D-Pro)-Gly turn and various residue mutations along the hairpin strands. They found that the inhibitor designed by Fasan *et al*. was the most flexible system (with 24% hairpin in solution by NMR Analysis of Molecular Flexibility in Solution, NAMFIS), and that the activity decreased for more constrained β-hairpins (with 61% hairpin in solution). This highlighted how an increased molecular flexibility correlates with an increase in inhibitory activity in this system.

Considerable progress has been made in developing and optimising these β-hairpin mimetics, but questions remain unanswered about the cell-permeating properties of these compounds as well as their inhibition specificity.[299]

4.4 Hypothesis

It was envisioned that the (D-Pro)-(L-Pro) turn in the inhibitor designed by Fasan *et al*. could be substituted by the designed β-turn mimic unit U-BTM (**28**, Figure 128). The sequence chosen for modification, **29**, was the same as the inhibitor with the best IC_{50} values, with all native amino acids. The sequence containing the (6-Cl)-Trp residue was not synthesised in this work due to the limited availability of this amino acid.

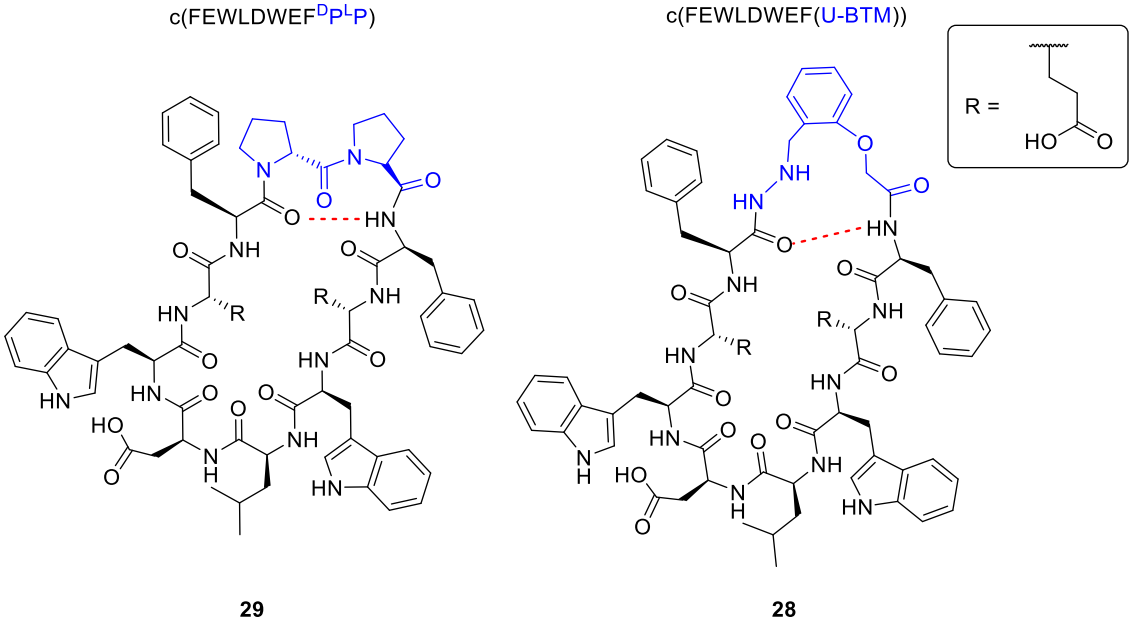


Figure 128. Structures of the control peptide 29 with a (D-Pro)-(L-Pro) turn (left), and the modified system containing the U-BTM unit instead, 28 (right). The *i* to *i+3* hydrogen bond is represented in red.

From the crystal structure of **27** bound to MDM2 available (Figure 129) and based on previous studies[300] as previously mentioned, it can be observed that the residues in the turn do not participate in binding. Therefore, the introduction of a β-turn mimic should not negatively affect the affinity of the inhibitor. A possible additional interaction that could be established by the mimic would be an aromatic interaction with 66Tyr, which is in close proximity to the L-Pro residue in the inhibitor **27**.

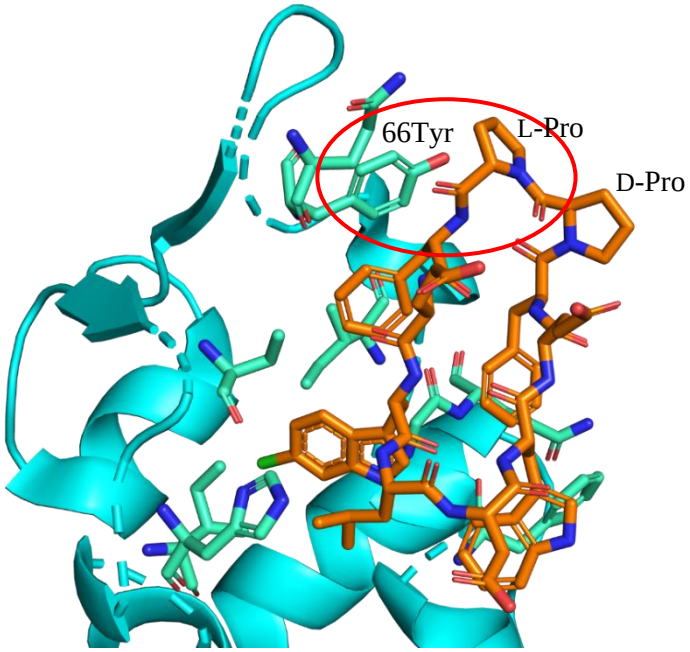


Figure 129. Close-up of the binding between MDM2 (in blue) and the inhibitor 27 (depicted in orange). PDB code 2AXI.

4.5 Aims

The aims defined for this chapter were as follows:

- Synthesis of the control sequence and of the mimic analogue, the latter containing a non-natural β-turn instead of a (D-Pro)-(L-Pro) turn.

- Evaluation of the ligation step as a means to cyclise the peptide.

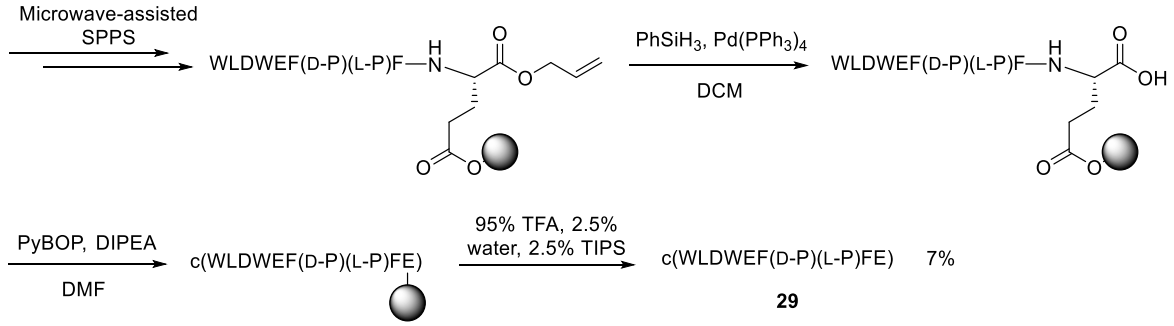
- Evaluation of the inhibition affinity of both sequences *via* Surface Plasmon Resonance (SPR) assays.

4.6 Synthesis of the control sequence, 29

In the papers by Fasan *et al.*[160] where these inhibitors are described, the synthesis described is performed on 2-Cl-trityl chloride resin using standard microwave-assisted SPPS followed by cleavage from the resin with 1% TFA in DCM, which keeps the protecting groups on the amino acid side chains. The cyclisation is performed in solution using HATU/HOAt activation, and then the peptide is deprotected through treatment with 95% TFA, 2.5% H_2O and 2.5% TIPS. In this work on-resin cyclisation was attempted instead, to obtain the desired

peptide directly upon cleavage and to prevent the loss of material that could occur if the cyclisation was performed in solution.

The control peptide **29** containing the (D-Pro)-(L-Pro) turn was synthesised on Wang resin preloaded with side chain-linked glutamic acid (Fmoc-Glu(Wang)-OAll, loading 0.44 mmol/g) on a 0.1 mmol scale (Scheme 29). The chain was elongated through standard microwave-assisted SPPS using DIC/Oxyma Pure activation, and then the allyl-protected *C*-terminus was deprotected using palladium(0) tetrakis(triphenylphosphine) and phenyl silane.[301,302] The head to tail on-resin cyclisation was performed using PyBOP and DIPEA[303] for 2.5 h and the peptide was then cleaved from the resin using standard conditions. After lyophilisation, the crude purity obtained was of 69%. Once purified *via* RP-HPLC, **29** was obtained in a 7% overall yield.



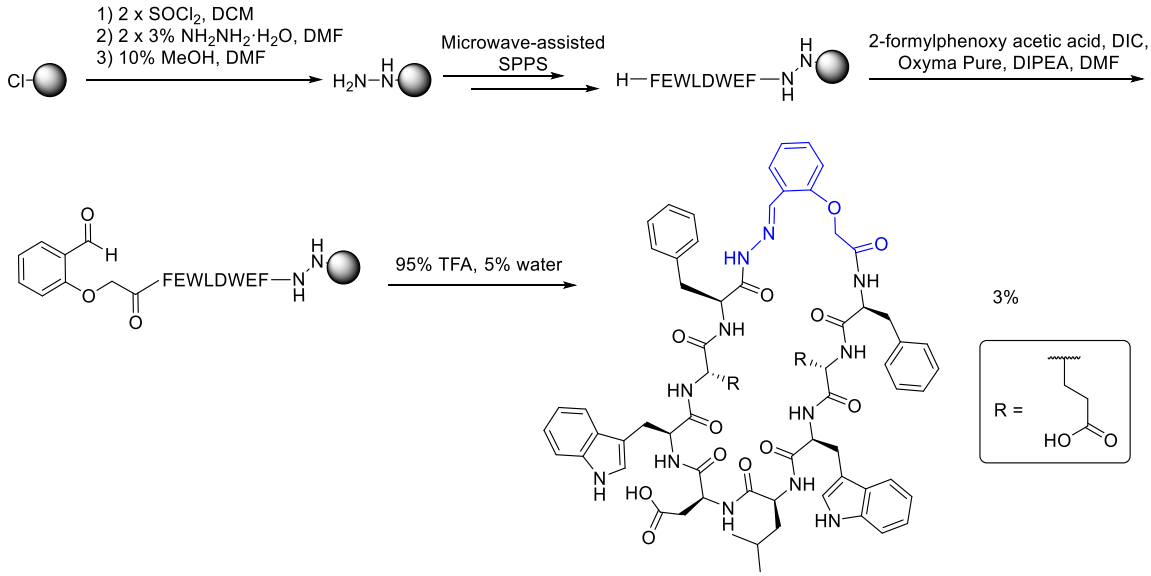
Scheme 29. Synthesis of the control peptide 29.

4.7 Synthesis of the mimic sequence, 28

The analogue containing the β-turn mimic, **28**, was synthesised on 2-Cl-trityl chloride resin (0.8 mmol/g loading) on a 0.1 mmol scale (Scheme 30). When hydrazine hydrate was used directly to modify the resin very low loadings were observed, which could be caused by hydrolysis of the resin. Therefore $SOCl_2$ was used to re-activate the resin, following conditions previously developed within the group. A solution of 3% $SOCl_2$ in DCM was added to the resin and shaken for 30 min. This step was repeated. A solution of 3% hydrazine hydrate in DMF was then added to the resin for 30 min, a step which was also repeated. After the hydrazine loading, a capping step was performed with 10% MeOH in DMF for 30 min. Peptide elongation to introduce the native amino acids was then performed through standard microwave-assisted SPPS, using DIC/Oxyma Pure activation.

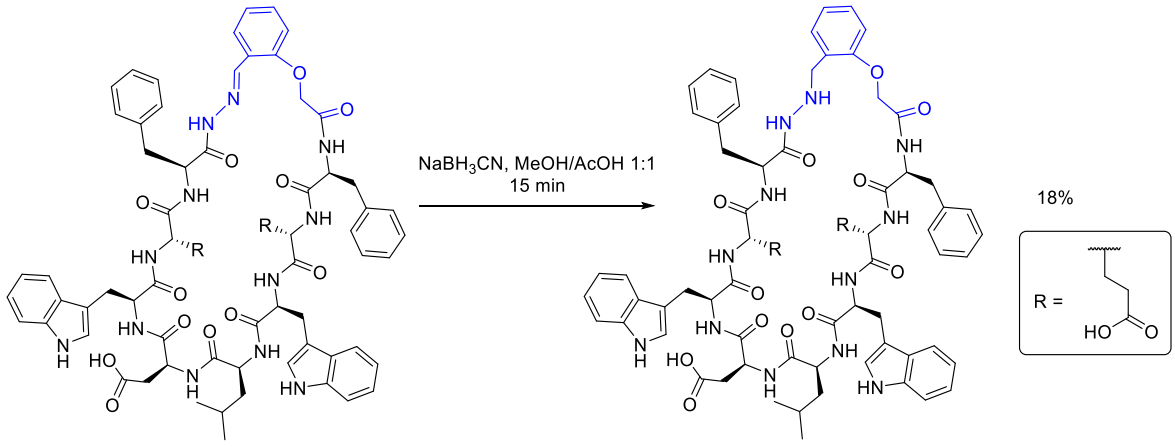
2-Formylphenoxyacetic acid was coupled at room temperature, using DIC/Oxyma Pure and DIPEA in DMF for 2.5 h. Cleavage from the resin was performed using TFA and water as

the scavenger, without the use of TIPS to prevent the reduction of the aldehyde. It was initially expected to obtain the linear aldehyde-containing peptide after cleavage, which would then be subjected to a reductive amination in order to cyclise the peptide in solution, but the cyclic hydrazone **30** was generated instead. It was therefore decided to isolate this intermediate and perform the reduction in solution after purification. **30** was obtained with a crude purity of 78%, and the overall yield after purification *via* RP-HPLC was of 3%.



Scheme 30. Synthesis of 30.

The reduction of the hydrazone to yield **28** (Scheme 31) was performed in MeOH/AcOH 1:1, using 10 equivalents of $NaBH_3CN$.



Scheme 31. Reduction step for the synthesis of 28.

HPLC monitoring showed that the reaction reached completion after 15 minutes, with full conversion to the hydrazide being observed (Figure 130). After purification, the mimic

analogue was obtained in an 18% yield. This confirmed the potential of the U-BTM unit and the ligation step to be used as an efficient peptide cyclisation strategy.

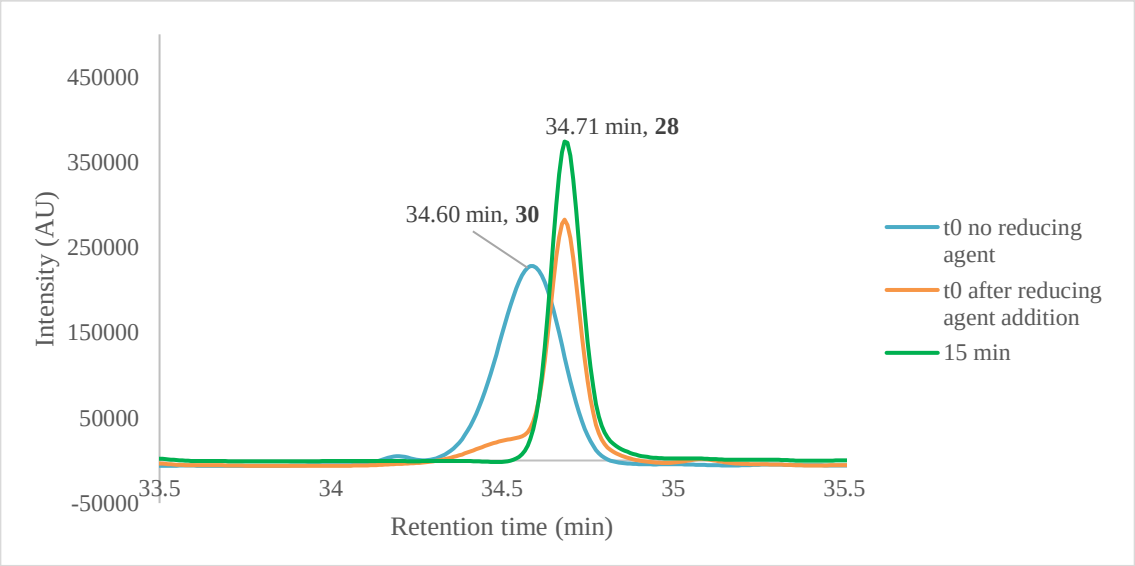


Figure 130. HPLC monitoring of the reduction to generate 28.

4.8 Surface Plasmon Resonance assays

Surface Plasmon Resonance (SPR) is one of the most used *in vitro* techniques to study protein-protein interactions. This method is based on the detection of changes in the refractive index of the medium in close proximity to a metal surface. The most common configuration for SPR (Figure 131, left) consists of a light source which shines polarised light through a prism, a chip with a metal surface where the light is reflected, and a detector. The intensity of the reflected light will vary depending on the angle of incidence of the polarised light on the metal chip (Figure 131, right). There is a particular angle of incidence at which the light will excite the electrons in the sensor chip (surface plasmons) and cause their resonance, which results in a loss of intensity of the light. The angle at which the maximum loss of intensity is observed is called the resonance angle or SPR-dip.[304]

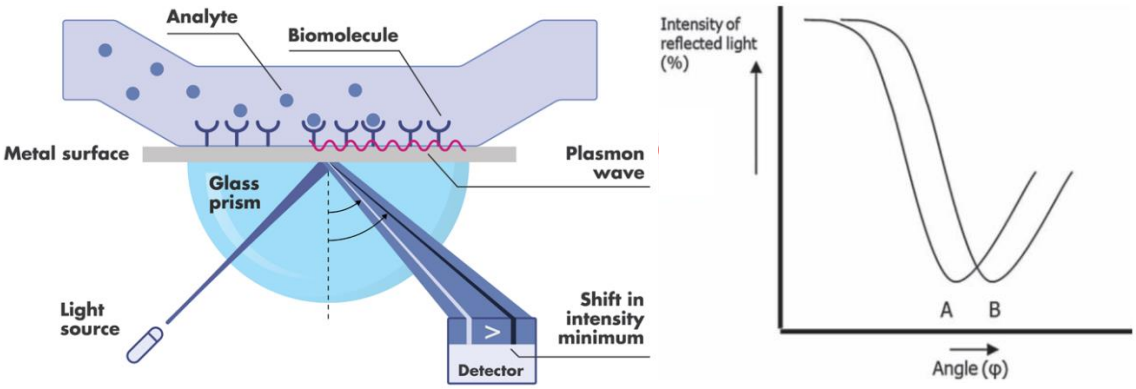


Figure 131. Schematic representation of the SPR assay set up (left)[305] and representation of the change in the intensity of the reflected light as a function of the angle of incidence (φ) (right). Reproduced with permission from ref.[304]

The SPR-dip changes as a function of the refractive index of the metal surface on the chip and its immediate environment. Thus, when a protein is accumulated on the surface the refractive index is altered, which causes a shift in the shape and location of the SPR-dip. This fluctuation can be used to gain real-time information about the sensor surface.

When an SPR assay is being set up, a ligand is immobilised on the surface of the chip and the binding partner or analyte is dissolved in the running buffer, which is flowed across the surface. Monitoring of the shift in the SPR-dip generates a sensogram, which allows the identification of the different stages of the assay (Figure 132).

In the first stage, and once the ligand has been immobilised on the chip surface, the surface is conditioned by flowing the running buffer over it. Once a stable baseline is achieved the analyte is injected, which initiates its association with the ligand. As more analyte binds to the ligand, the SPR response increases. At this stage the association rate constant (k_a) can be calculated. This continues until equilibrium is reached, where the association and dissociation rates become equal. The last stage of the assay is the injection of buffer, which causes the dissociation of the analyte. This allows for the calculation of the dissociation rate constant (k_d) and leaves the surface ready for a new injection cycle.[304,306]

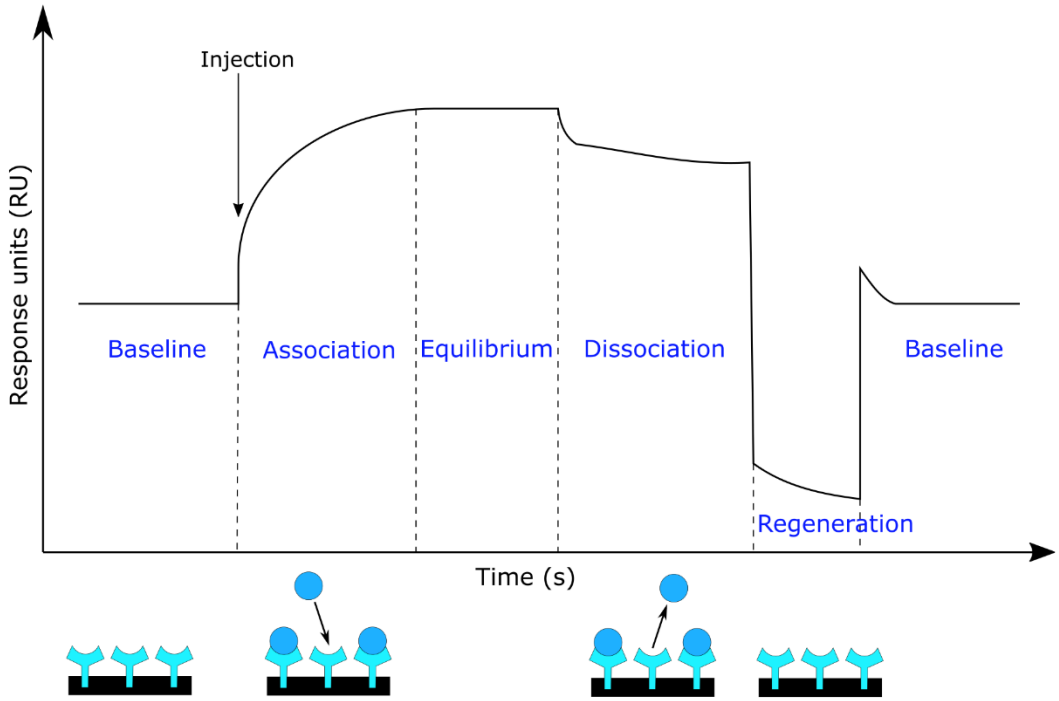


Figure 132. Schematic representation of the sensogram of an SPR assay, showing the different phases of the analysis.

When the system is in equilibrium, the equilibrium dissociation constant (K_D) can be calculated, which gives information on the affinity of the biomolecular interaction. It can be expressed by the concentrations of the species involved as well as the rate constants (Equation 17). A high K_D value indicates a low stability of the analyte-ligand complex, which means that there is a small binding affinity between partners.[304]

$$A + L \underset{k_d}{\overset{k_a}{\rightleftharpoons}} AL \qquad K_D = \frac{[A][L]}{[AL]} = \frac{k_d}{k_a}$$

Equation 17. Calculation of the equilibrium dissociation constant, where [A] is the concentration of the analyte, [L] the concentration of the ligand and [AL] the concentration of the complex. The units for K_D are mol L^{-1}.

4.9 Evaluation of the binding affinity using SPR

SPR assays on the control and mimic peptides were performed by Dr Mads Gabrielsen, working in Prof Danny Huang's group at the Beatson Institute. The MDM2 p53-binding domain (residues 1 to 109) was tagged with a glutathione S-transferase (GST) tag, and it was then bound onto a GST-antibody which was coupled on the chip surface. The peptide in study was then diluted into the running buffer and flowed across the chip at different concentrations. Seven concentrations were used, increasing from 0 to 50 μM and *vice versa*, and a control with only GST was also run. All repeats were reproducible. The data points

were then fitted using a single-site binding model in PRISM, which generated a binding curve from where the K_D was estimated.

The sensogram and fitted data for the control sequence **29** are presented in Figure 133:

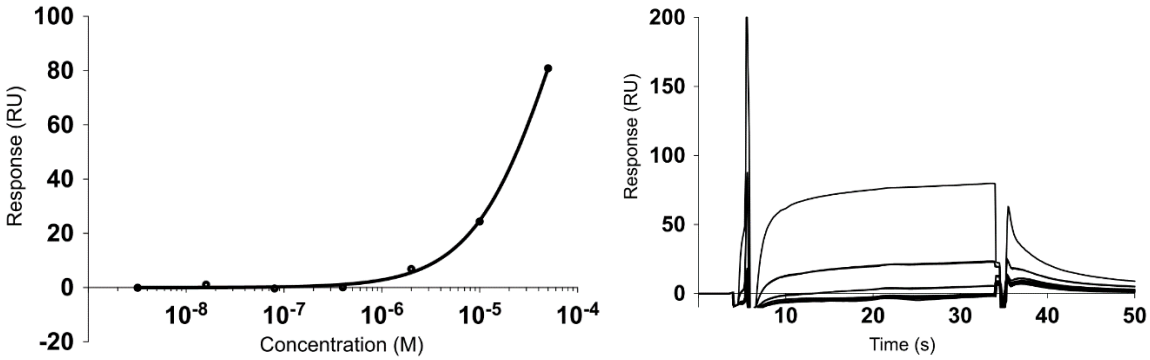


Figure 133. Left: binding curve for 29. Right: sensogram for 29. The different lines in the sensogram correspond to all the concentrations tested.

The sensogram and fitted data for the mimic **28** are presented in Figure 134. As it can be observed for both peptides, an increase in response units takes place as the peptide binds to the anchored protein with increasing concentrations. Looking at the binding curves, in the case of **28** the same concentration gives rise to a higher response in comparison to **29**, which indicates higher binding.

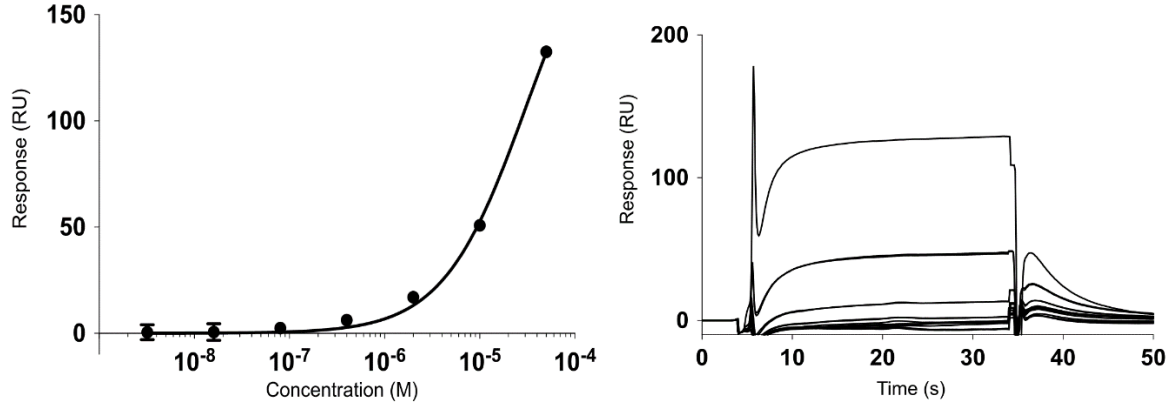


Figure 134. Left: binding curve for 28. Right: sensogram for 28. The different lines in the sensogram correspond to all the concentrations tested. The error bars presented correspond to the standard deviation of the mean between two repeats for each concentration.

The K_D obtained for **29** was 65.6 ± 6.3 μM, and that of **28** was 30.3 ± 3.5 μM (the uncertainties for these values arise from the standard deviation of repeats). This not only showed maintenance of the affinity of the inhibitor when the turn was substituted by the

unnatural U-BTM unit, but a two-fold improvement in binding for the mimic peptide. The previously mentioned study by Danelius *et al.*[300] showed that an increased molecular flexibility in β-hairpin-based inhibitors like **29** correlated to an increased p53/MDM2 inhibitory activity. The increased affinity of **28** could be due to the higher flexibility of the turn mimic in comparison to the (D-Pro)-(L-Pro) turn, which is conferred by the sp^3-hybridised nitrogen and carbon atoms on the hydrazide linkage as well as the CH_2 unit on the phenoxyacetyl-like centre.

Another reason for the increased affinity could be the establishment of additional interactions between the U-BTM unit and the MDM2 binding pocket. As mentioned previously (see Figure 129), the crystal structure of the previously reported inhibitor bound to MDM2 shows that the residue 66Tyr in MDM2 is in close proximity to the (D-Pro)-(L-Pro) turn of the inhibitor. When introducing the U-BTM unit an aromatic interaction could take place with that residue, possibly with an edge-to-face geometry, which would improve the affinity to the protein. Another possibility is the presence of an additional hydrogen bonding interaction between the hydrazide unit in U-BTM and 59Gln or 62Met in MDM2, which in the crystal structure appear to be close to the (D-Pro)-(L-Pro) turn.

The binding for the control peptide used in this work, **29**, has been previously evaluated by Fasan *et al*. although only as an IC_{50} value. In contrast, the binding of the sequence containing a (6-Cl)-Trp residue has been quantified both as an IC_{50} and as a K_D (Table 15). In comparison to the K_D obtained by Danelius *et al*., the value obtained in this work for **29** is two orders of magnitude weaker, a difference most likely caused by the change to a native Trp residue.

Table 15. Binding affinities found for previously reported inhibitors as well as the systems studied in this work.

	Sequence	K_D	IC_{50}
Fasan *et al.*[160]	c[**W**LDWEF**(D-Pro)-(L-Pro)**FE]		0.53 μM
	c[**(6-Cl)-W**LDWEF**(D-Pro)-(L-Pro)**FE]		0.14 μM
Danelius *et al.*[300]	c[**(6-Cl)-W**LDWEF**(D-Pro)-(L-Pro)**FE]	127 nM	2.86 μM
This work	**29**, c[**W**LDWEF**(D-Pro)-(L-Pro)**FE]	65.6 μM	
	28, c[**W**LDWEF**(U-BTM)**FE]	30.3 μM	

After careful evaluation of the SPR assays set up and samples that were used, our collaborators at the Beatson Institute found that the MDM2 construct that was employed was unstable due to the sequence being too short. This was affecting the binding affinities that were measured, and therefore the experiments will need to be repeated. These experiments are currently being performed, using a longer protein construct to ensure its stability. This will provide K_D values closer to those previously reported and will ensure their reproducibility.

4.10 Conclusions and future work

A novel inhibitor of the MDM2/p53 interaction has been designed, which contains a non-proteinogenic β-turn mimic instead of the (D-Pro)-(L-Pro) turn present in the previous inhibitors reported.[160] The (D-Pro)-(L-Pro) control and the mimic peptide have been synthesised, and the ligation protocol has been validated as a strategy for its use on cyclic systems. SPR assays have been run by collaborators, which not only showed that inhibitory affinity was maintained upon introduction of the unnatural turn unit, but a two-fold improvement was observed for the mimic with respect to the control. This confirmed the potential of the designed U-BTM mimic to be used within biologically relevant targets. This system is the first β-hairpin inhibitor of the MDM2/p53 interaction that contains a completely non-natural β-turn mimic, and one of the first ones in the inhibition of PPIs more generally.[136]

The SPR assays will be repeated by Prof Danny Huang's group with a more stable MDM2 construct, in order to obtain more accurate K_D values. It would also be interesting to evaluate whether these cyclic peptides are cell-penetrant, and if so to quantify their activity in this context. Cell-based assays will be performed by Prof Danny Huang's group, which will quantify the activity of the peptides described in this chapter within cells.

If the peptides are not cell-penetrant, their overall charge could be altered to achieve this. The charge of **28** at physiological pH is -2, as the U-BTM unit introduces a positive charge at the hydrazinium nitrogen, but usually a positive charge is required for peptides to enter cells. To achieve this, the Asp and Glu residues could be exchanged for Lys, Arg or His. Another strategy to favour cell penetration would be the modification of the U-BTM unit: the aromatic ring could be functionalised further to introduce charged groups or motifs like a poly-Arg sequence, and the flexibility of the hairpin could also be altered. If necessary, the ligation junction could also be changed to an oxime linkage for example, in order to remove the charge.

Future work could also involve the synthesis of the mimic sequence with the (6-Cl)-Trp residue instead of native Trp, as previous work by Fasan *et al.*[160] has shown that the inhibitory activity is considerably improved when using this non-natural amino acid.

5 β-Turn mimics through oxime ligation

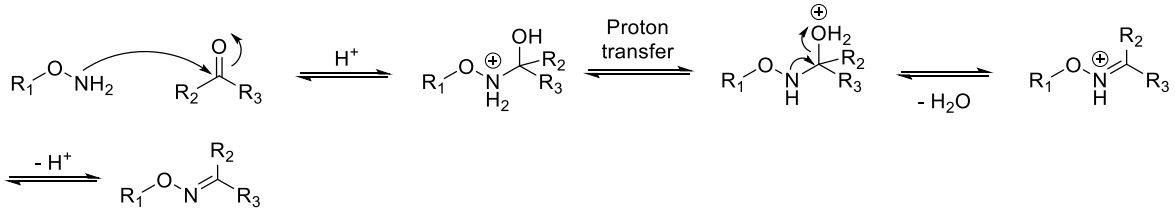
The work presented in this chapter was performed in collaboration with Dr Drew Thomson, who performed the molecular dynamics simulations.

5.1 Oxime ligation

Oxime ligations are chemoselective reactions that have attracted considerable interest for a wide range of applications in peptide and protein science. Although oxime formation is slow, this conjugation strategy provides a junction with a higher stability against hydrolysis at physiological pH in comparison to hydrazones.[307] These properties have made oxime ligation an attractive strategy for the synthesis of cyclic peptides,[308] proteins,[309] as staples to promote helicity[202] and in bioconjugation.[310]

5.1.1 Synthesis of oximes

Oximes are the product of a condensation reaction between an aminooxy and a carbonyl fragment (Scheme 32). Oxime ligations are relatively slow at physiological pH, but they can be accelerated if performed at a pH of 4-5 or if nucleophilic catalysts are employed. The nature of the aminooxy[311] and the carbonyl fragment also has a considerable effect on the reaction rate as well as on the stability of the ligated product.[193] Aldehydes are usually more reactive than ketones due to steric effects, which is why aldoximes have been more studied than ketoximes.[310] Aliphatic aldehydes react more rapidly than aromatic ones,[312,313] but the latter provide a higher stability of the oxime product.[314] In contrast to aldehydes, aliphatic and aromatic ketones have been found to react at similar rates,[312] and their reactivity depends on the stability of the ketonium carbocation formed under acidic conditions.[315]



Scheme 32. Mechanism of oxime ligation.

The fact that ketoximes have an additional substituent at the ligation junction could allow for the tuning of the properties of the oxime bond as well as providing a better mimicry of native peptide sequences[316] through the inclusion of groups that would mimic amino acid

side chains. Although this is an advantage over aldoximes, most applications of the oxime bond utilise aldoximes instead of ketoximes due to the slow reactivity of the latter. Aliphatic and phenyl-substituted ketoximes were used by Guthrie *et al.*[316] to mimic parallel β-sheets, and these linkages have also been employed in the synthesis of cyclic peptides,[317] but the *E*- and *Z*-oxime isomers have not been identified in most of these studies. Lamping *et al.*[318] characterised the ketoxime stereoisomers in oxime-ligated hexapeptides which contained two β-turns, one with a Pro residue and another with the oxime linkage (Figure 135). Although four conformations were expected due to the *E*/*Z*-oxime isomerism and the *cis/trans* isomerism of Pro, they found that two conformations were adopted (arising from the Pro *cis/trans* isomers, Figure 135) and that the *E*-ketoxime was the only stereoisomer present. This could indicate a higher stability of the *E*-stereoisomer in this system. However, the isomerism and stability of the ketoxime stereoisomers and their impact on peptide structure and conformation has been scarcely explored in the literature.

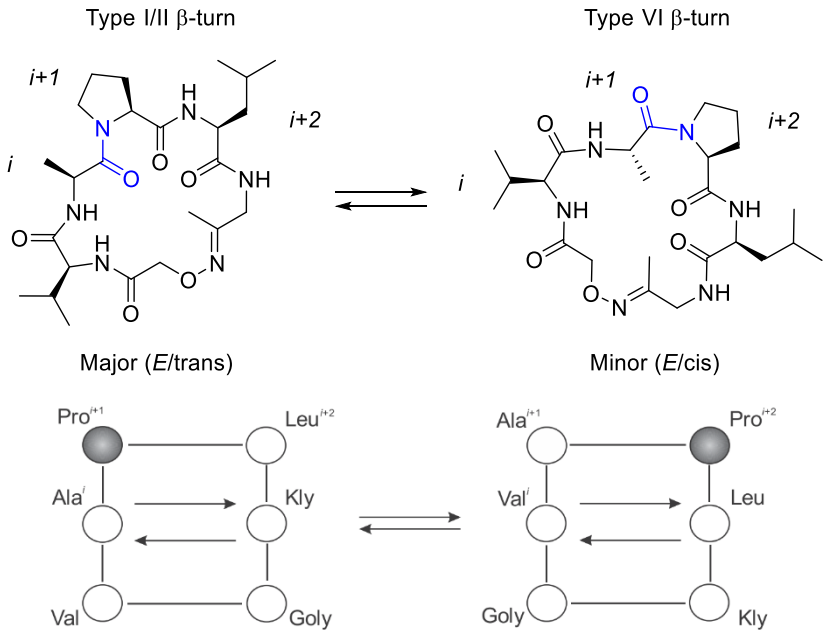


Figure 135. Conformations adopted by the hexapeptides synthesised by Lamping *et al.* Reproduced with permission from ref.[318]

Many strategies are available to introduce the required functionalities into the ligation partners. The aminooxy moiety is usually introduced by means of an unnatural amino acid like aminooxyacetic acid (Aoa, see Figure 136). Aoa can be protected with a propargyloxycarbonyl (Proc)[319] or Boc protecting group[309] to prevent its reaction with acetone or ambient traces of aldehydes like formaldehyde. The carbonyl fragment can be introduced into the peptide sequence by coupling ketoacids like levulinic or pyruvic acid (Figure 136), as well as by using modified amino acid side chains.[307] Aldehyde masking

strategies can be used to prevent undesired side reactions, for example by introducing an unnatural amino acid with a thiazolidine[320] or a diol moiety which is later reacted to unmask the aldehyde.[321] In addition, oximes can also be obtained by oxidative coupling processes in which a carbonyl moiety is not required.[316]

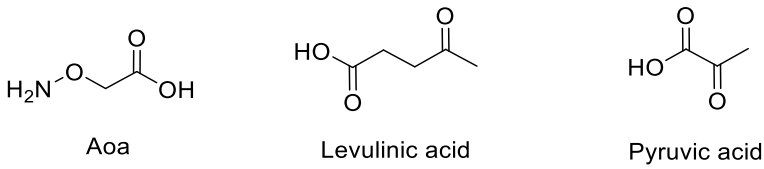
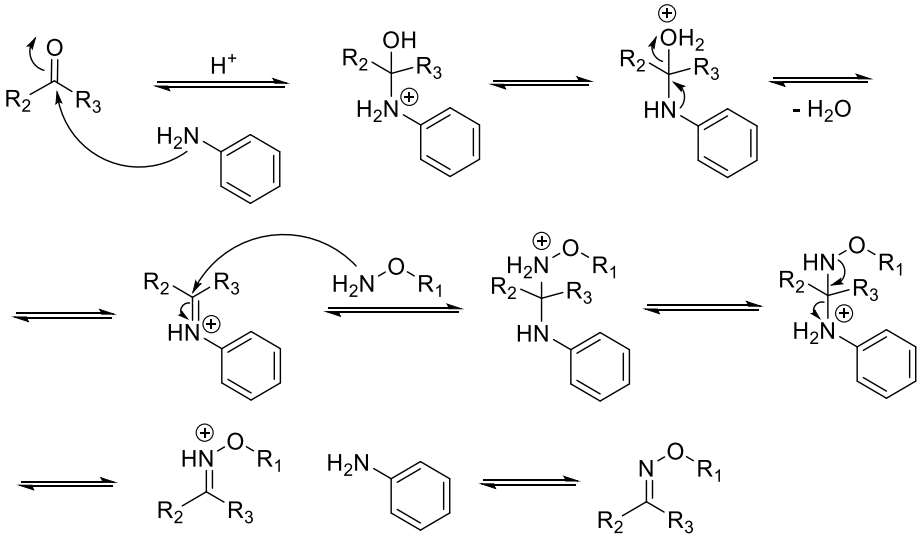


Figure 136. Building blocks used to generate the oxime ligation partners.

5.1.1.1 Nucleophilic catalysis

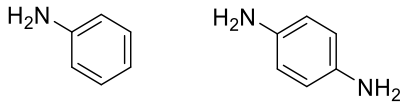
Oxime ligations are fastest at a pH of 4.5 and at above millimolar-range concentrations,[194] but in general these reactions are still slow, taking days to reach completion. In order to improve the reaction rates, especially at physiological pH, multiple catalysts have been developed, which are referred to as nucleophilic catalysts. One of the first catalysts introduced and one of the most used is aniline (Scheme 33): Dirksen *et al.*[195] found that the reaction rate of an oxime ligation between two unprotected peptides (at 1 mM concentration for each peptide) could be accelerated 40-fold at neutral pH (in 0.1 M sodium phosphate buffer) using 10 mM aniline, and 400-fold at a pH of 4.5 (in 0.1 M ammonium acetate buffer) using 100 mM aniline.



Scheme 33. Mechanism of aniline catalysis in an oxime ligation.

Other catalysts have also been explored, which utilise substitutions on the aniline ring to enhance its catalytic efficiency.[322] To this end, Wendeler *et al.*[323] studied substituted anilines

and found that catalysts containing an electron donating amino, hydroxy or methoxy group at the *para*- position of the aniline ring were all superior to aniline at neutral pH. In contrast, the *ortho*- and *meta*- isomers of these compounds were much less effective. The most efficient catalyst found was *p*-phenylenediamine (*p*-PDA, Figure 137), which was superior to aniline in the pH range of 4 to 7: a rate enhancement of 4 to 19-fold was observed in comparison to the aniline-catalysed ligations, with the greatest difference occurring at higher pH values. A 10 mM concentration of the catalyst was used in the ligations (with a concentration of 90 μM for the aldehyde-containing partner and 450 μM for the *N*-alkoxy amine-containing partner, in sodium citrate buffer 100 mM), but when *p*-PDA was tested at a 2 mM concentration the same rate enhancement was observed.



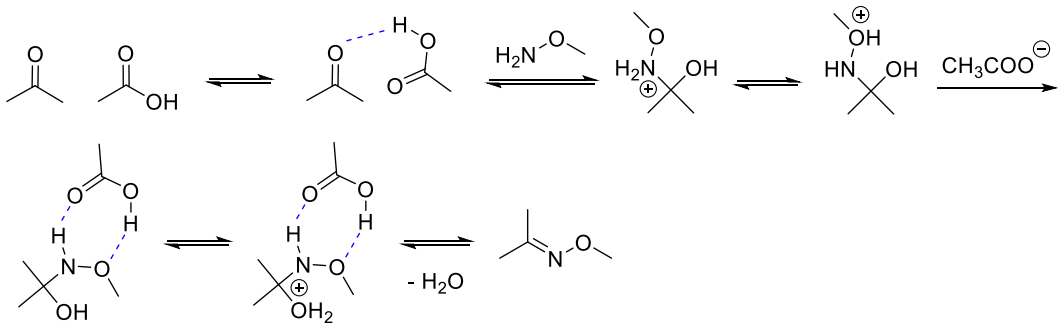
Aniline *p*-Phenylenediamine

Figure 137. Structures of nucleophilic catalysts aniline and *p*-PDA.

Although aniline and *p*-PDA are suitable catalysts for aldehyde substrates, their efficiency towards ketone substrates is limited.[324] This has prompted the exploration of other strategies to accelerate the reaction rates of ketoxime formation. The catalysis of transimination by acid and base buffers was first recognised by Jencks in 1977,[325,326] and acetic acid is commonly used in oxime ligations as it has been shown to improve the solubility of the substrates as well as accelerate the reaction rates. Chelushkin *et al*.[307] reported that using neat acetic acid and peptide concentrations over 2.5 mM afforded over 95% conversion to the ligated product in 1.5-2 h.

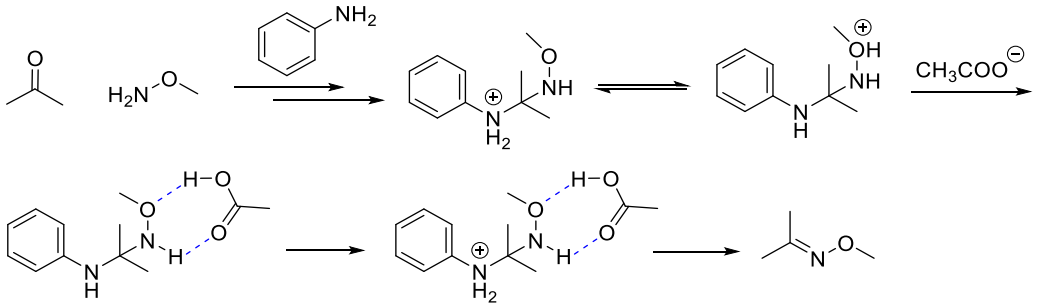
Wang *et al*.[324] found that using 1 M acetic acid during a ketoxime ligation caused a reaction rate increase of 4.3-fold in comparison to the uncatalysed reaction, and that using 1 M acetic acid as a co-catalyst with aniline (100 mM) increased the reaction rate 11-fold (these ligations were performed in citrate buffer 1 mM, with 10 mM ketone component and 1 mM *N*-alkoxy amine component). The authors suggested that this acceleration was caused by the establishment of hydrogen bonds between the ketone substrate and the acid, as well as between key intermediates and the acid. It was further suggested that the presence of these hydrogen bonds favoured the formation of the oxime product. Attenuated total reflectance Fourier transform infrared spectroscopy showed that when using acetic acid in the absence of aniline (Scheme 34), hydrogen bonds formed between the catalyst and the ketone

substrate, which would enhance the electrophilicity of the carbonyl centre and therefore favour the nucleophilic attack of the *N*-alkoxy amine partner. These experiments also showed that hydrogen bonds took place between the catalyst and the hemiaminal intermediate, which would favour the key dehydration step.



Scheme 34. Acetate-catalysed ketoxime formation. Hydrogen bonds are represented in blue.

When acetic acid is used as a co-catalyst with aniline, the increase in the reaction rate is even higher due to the formation of hydrogen bonds with the anilinium intermediate (Scheme 35).



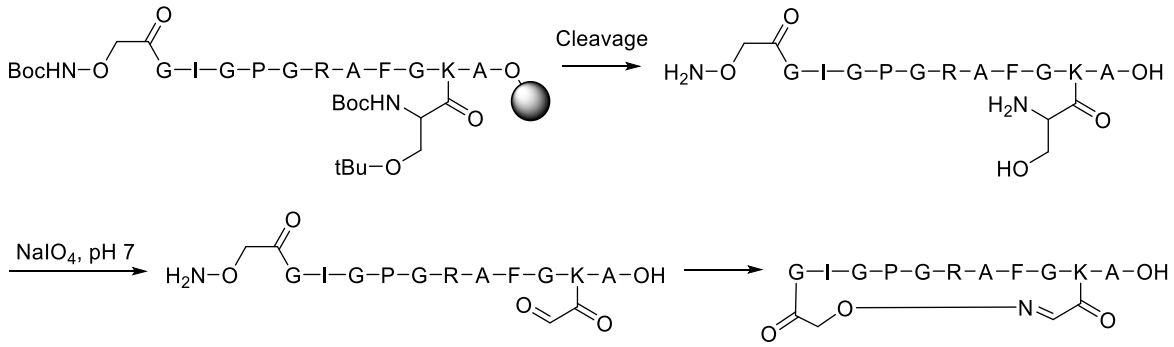
Scheme 35. Acetate and aniline-catalysed ketoxime formation. Hydrogen bonds are represented in blue.

Other catalysts have been explored to improve the reaction rates of ketoxime formation. For example, Larsen *et al.*[312] found that for the reaction between phenylhydrazine and acetophenone, 1 mM 2-(aminomethyl)-5-methoxybenzimidazole provided a 5.4-fold increase of the reaction rate in comparison to the noncatalysed ligation.

5.2 Applications of oximes

The fact that oxime ligations take place under mild conditions, are stable to hydrolysis, and the *N*-alkoxy amine and carbonyl moieties are easy to access makes oximes a versatile linkage that has found a wide range of applications.[308]

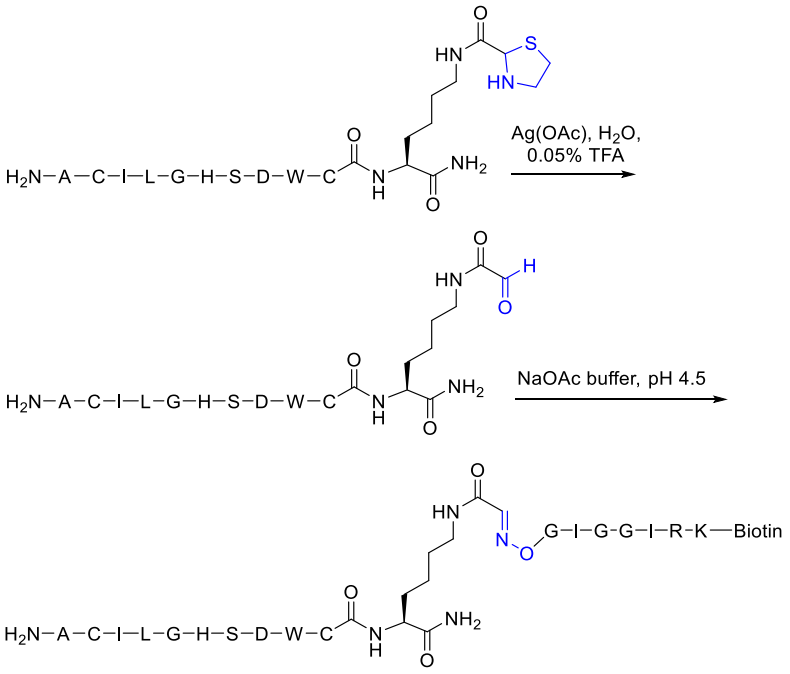
Oximes have been used in the synthesis of cyclic peptides, both employing a head to tail strategy[318] as well as cyclising through residue side chains.[198] An example of the latter is the peptide synthesised by Pallin *et al.*[327] (Scheme 36), who used a modified Lys residue to generate a glyoxylyl group on the *C*-terminus and introduced the *N*-terminal *N*-alkoxy amine through the coupling of aminooxyacetic acid. Generation of the aldehyde from the glyoxylyl moiety using $NaIO_4$ at pH 7 initiated the ligation process, which afforded the cyclised peptide in a 90% yield.



Scheme 36. Oxime-cyclised peptide synthesised by Pallin *et al.*[327]

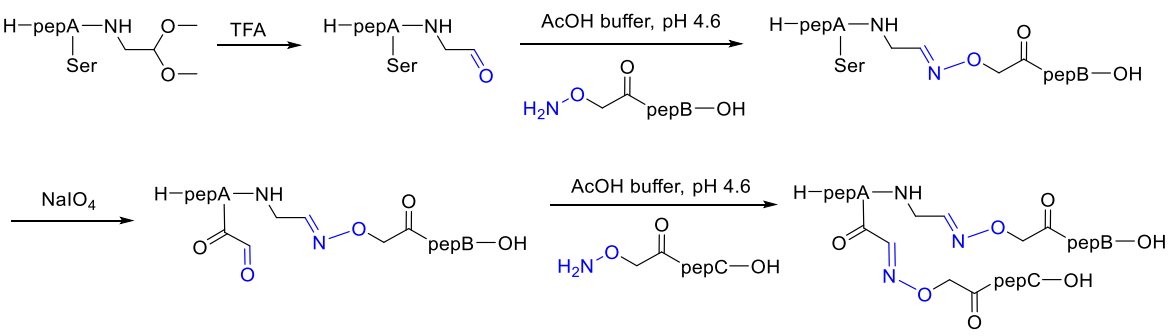
The stabilisation of α-helices is another area where oxime linkages have been utilised. Haney *et al.*[202] developed unnatural amino acids containing an aminooxy and a 1,2-aminoalcohol group and introduced them within α-helical backbones with *i* to *i+3* or *i* to *i+4* spacings. After treatment with sodium periodate, the aldehyde was generated on the 1,2-aminoalcohol residue and the oxime crosslink was formed at pH 7. It was found that the crosslink stabilised the helical conformation, with the *i* to *i+4* spacing providing a better folding than *i* to *i+3*, and that both *E*- and *Z*-oxime stereoisomers took place and were interconverting at neutral pH. The authors have subsequently used this strategy in the stabilisation of dimeric coiled-coils[201] and in the synthesis of α-helical peptides stapled with small molecule linkers, which can be used to generate libraries of cyclic peptides.[328]

One of the most interesting uses of oxime ligation is in the field of bioconjugation, due to its biorthogonality with the native amino acid side chains, the variety of strategies available for carbonyl incorporation and the possibilities for catalysis.[329] Bi *et al.*[320] used this strategy to site-specifically label a peptide and a recombinant protein (Scheme 37). To accomplish this, the peptide and protein were synthesised with a *C*-terminal thiazolidine-containing amino acid. After unmasking the aldehyde with silver acetate, an oxime ligation was performed with the aminooxy-containing partner (a biotinylated peptide) in sodium acetate buffer at a pH of 4.5.



Scheme 37. Strategy used by Bi *et al*.[320] to synthesise the labelled peptide analogue through oxime ligation.

Successive oxime ligations can be used in the stepwise synthesis of bioconjugates by sequential unmasking of the aldehyde or aminooxy moieties.[308] For example, Lelièvre *et al*.[330] employed this approach to synthesise a branched peptide (Scheme 38). To this end, the two carbonyl-containing peptide partner was synthesised with one aldehyde masked as an acetal, and the other masked as a 2-amino alcohol on a Lys residue acylated with Ser. The acetal-masked aldehyde was deprotected with TFA and the first oxime ligation was performed in an acetate buffer at pH 4.6 with an aminooxy-containing peptide partner. The aminooxy unit was introduced by coupling aminooxyacetic acid on the *N*-terminus. Once the first ligation was complete, the second aldehyde was unmasked by means of $NaIO_4$ oxidation and the final ligation step was performed with another aminooxy-containing peptide with a different peptide sequence.



Scheme 38. Synthetic route towards the branched peptide by Lelièvre *et al*.[330]

5.2.1 Oximes in the mimicry of peptide secondary structure

Vanderesse *et al.*[146] studied oxime moieties as peptide bond surrogates within dipeptides of the sequence Piv-Pro-[CH=NO]-Gly-NH^{i}Pr, and found that the *E*-oxime stereoisomer adopted an extended, flexible conformation, whereas the *Z*- stereoisomer formed a turn-like conformation.

Ketoxime linkages were used by Guthrie *et al.*[316] to mimic parallel β-sheets (Figure 138). The ligation was performed at pH 7 and under an oxygen atmosphere, which afforded three different ketoxime-bound conjugates, two of which contained a phenyl-functionalised junction. These structures could give rise to turn-like conformations, but their folding was not characterised in this work.

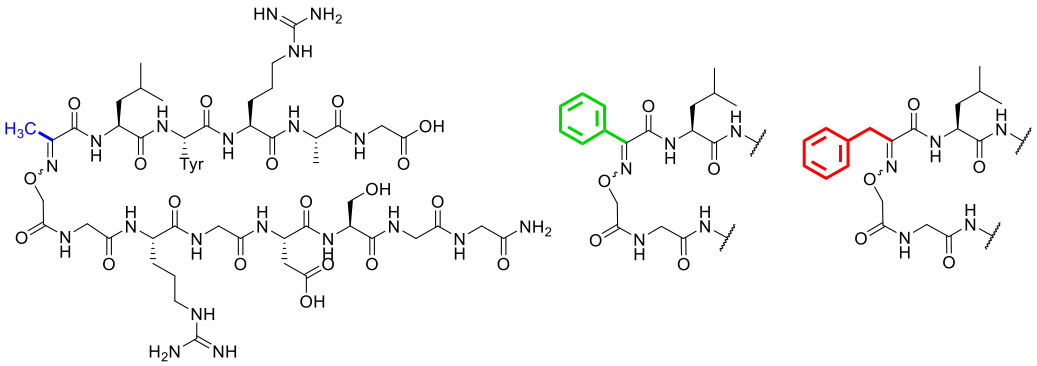
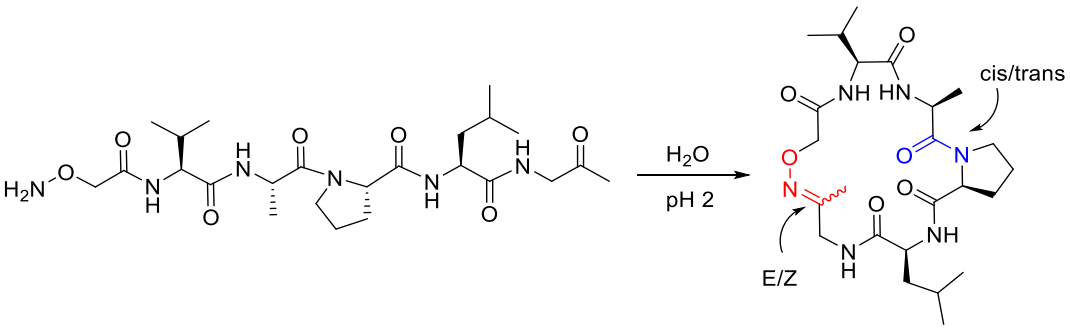


Figure 138. Ketoxime-linked parallel β-sheets synthesised by Guthrie *et al.*[316]

A ketoxime was employed by Lamping *et al.*[318] to cyclise a hexapeptide from head to tail (Scheme 39). This afforded a macrocycle that adopted two distinct conformations, the minor one containing a *cis* and the major a *trans* Pro. In both cases the ketoxime was present as the *E*- stereoisomer, and the *Z*-oxime was not observed. Importantly, the main conformer had a well-defined conformation that was equivalent to that of native double β-turn cyclic hexapeptides, which confirmed the potential of oximes to form part of a β-turn structure.



Scheme 39. Oxime ligation to generate the cyclic hexapeptide synthesised by Lamping *et al*.[318]

Similar to oximes, aminooxy groups have been used as part of peptidomimetics (named α- and β-aminooxypeptides, Figure 139). These moieties provide more rigid structures than natural analogues due to the repulsion between the lone pairs in the heteroatoms, which facilitates the establishment of hydrogen bonds between adjacent residues. These hydrogen-bonded structures adopt reverse turn conformations which are named N-O turns (Figure 139).[331]

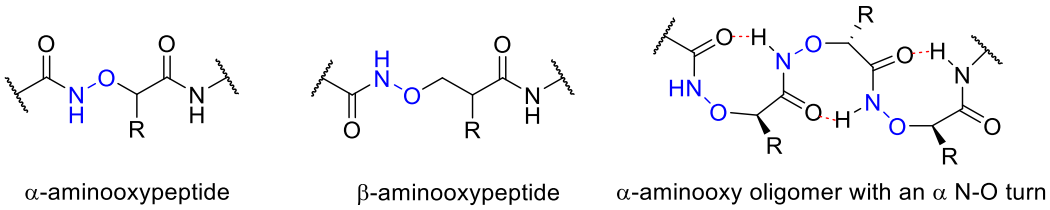


Figure 139. α- And β-aminooxypeptides, and structure of the N-O turns adopted by the former (the hydrogen bonds are represented in red).

5.3 Chapter aims

Oxime formation is an attractive ligation strategy that provides a junction with higher hydrolytic stability than hydrazones. Although it has found a variety of applications in bioconjugation, the synthesis of cyclic peptides and the stabilisation of α-helical structures, only a few examples exist where oximes have been employed to mimic elements of protein secondary structure.

In this chapter novel β-turn mimic scaffolds would be designed, which would be accessed through an oxime ligation step between two peptide fragments, one containing a carbonyl and the other an *N*-alkoxy amine moiety. Two types of scaffolds were to be explored: firstly, an aromatic-ring based linker, and secondly, an aliphatic one, the comparison of which would allow for the evaluation of the level of preorganisation required for a β-turn geometry

to be adopted. The mimics would be implemented within the control β-hairpin **TrpZip1**, introduced in Chapter 3.

The aims of this chapter were the following:

- To design the oxime-based β-turn mimics to achieve a suitable reverse turn geometry.

- To develop suitable oxime ligation conditions and synthesise the β-hairpin conjugates.

- To evaluate the conformation and thermal stability of the conjugates through CD experiments and ascertain whether the β-hairpin structure is maintained.

5.4 *Ortho*-substituted oxime mimic

The first target designed consisted of an *ortho*-substituted aromatic ring (Figure 140), which was expected to provide a suitable geometry for the mimicry of the β-turn (as it was confirmed with the β-turn mimics explored previously in Chapter 3).

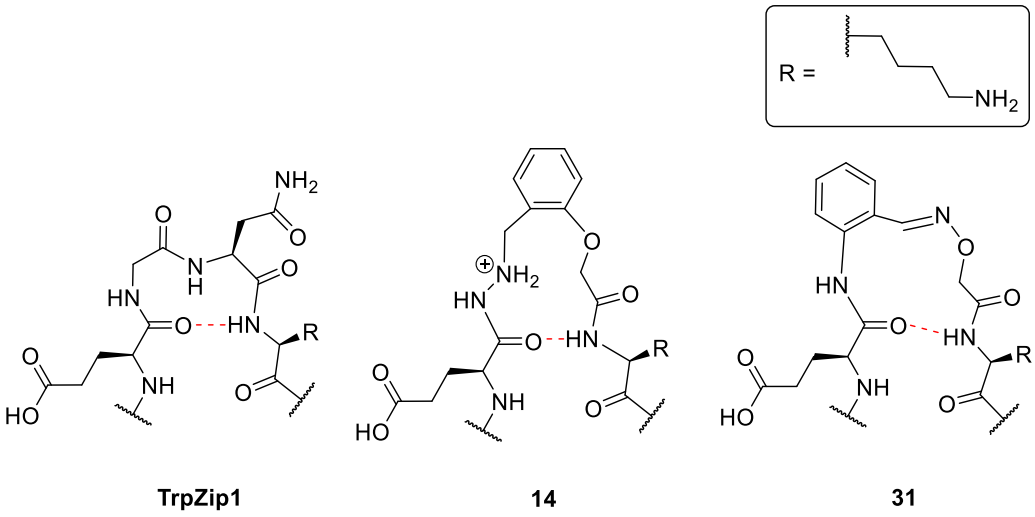
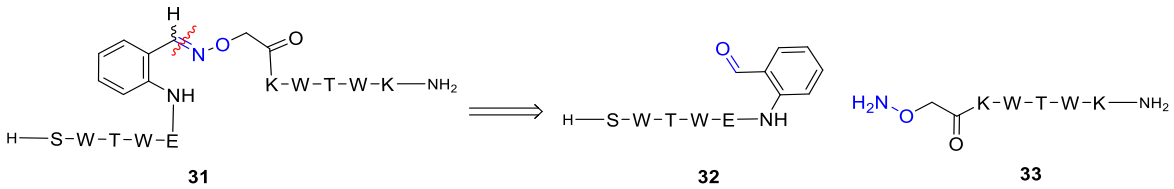


Figure 140. Comparison between the GN turn of TrpZip1 (left), the hydrazide-linked turn of 14 (middle) and the oxime-ligated turn in mimic 31 (right).

This structure would be analogous to that of the hydrazide-linked **14**, **(*S*)-15** and **(*R*)-15** systems (see Chapter 3), but with the functionality in the peptide fragments being inverted. In this case, the carbonyl unit would be coupled on the *N*-terminal peptide, **32**, whereas the *C*-terminal peptide **33** would provide the *N*-alkoxy amine functionality. The target β-hairpin

would be accessed through a ligation step between the fragments that would generate an oxime junction (Scheme 40).



Scheme 40. Oxime-linked target, 31, and the two peptide substrates for the ligation.

Energy minimisation experiments were performed for the two stereoisomers in the mimic system to obtain preliminary information on whether the β-hairpin and turn conformation would be maintained. The models showed that both conjugates had a distance of 3 Å between the oxygen in the carbonyl of the residue preceding the mimic and the following nitrogen. Although the system with the *E*-oxime seemed to have a less ideal hydrogen bond geometry (Figure 141), both systems appeared to maintain the β-hairpin conformation.

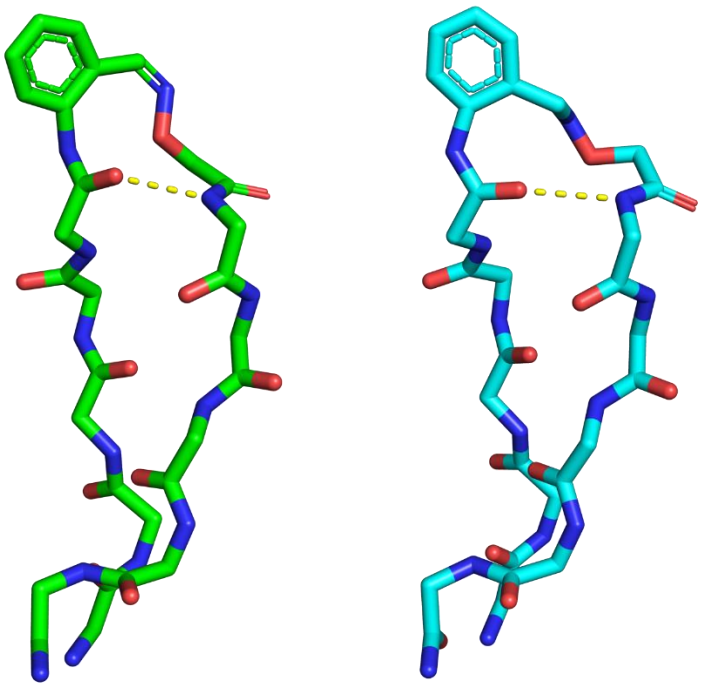
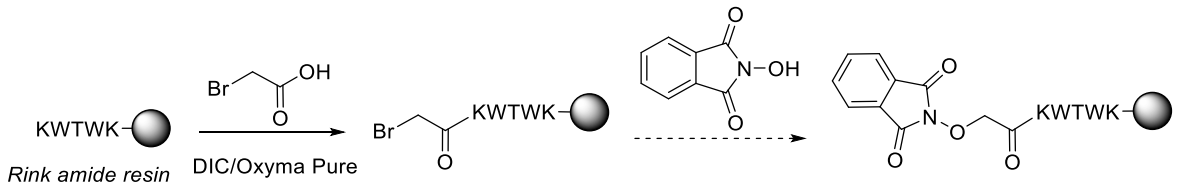


Figure 141. Energy minimisation models obtained for *Z*-31 (left) and *E*-31 (right), showing the backbone atoms only and the *i* to *i+3* like hydrogen bond in yellow.

5.4.1 Synthesis of the *C*-terminal fragment 33

The synthesis of **33** started with microwave-assisted SPPS using DIC/Oxyma Pure activation to couple the native amino acids onto Rink Amide resin (0.51 mmol/g loading) in a 0.1 mmol scale. It was initially attempted to couple bromoacetic acid (5 equiv.) on the peptide using DIC (4.5 equiv.) and Oxyma Pure (4.5 equiv.) at room temperature for 2.5 h (Scheme 41).



Scheme 41. Initial synthetic route for the synthesis of 33.

A test cleavage was performed on a small amount of resin after the coupling using TFA, H_2O and TIPS (95/2.5/2.5%), which was incubated for 45 min. After evaporation of the TFA under a stream of nitrogen, the sample was re-dissolved in H_2O/MeCN and analysed by LCMS, which showed that the Oxyma adduct (Figure 142) was generated at the α-carbon.

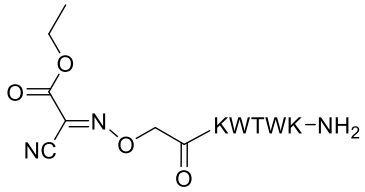
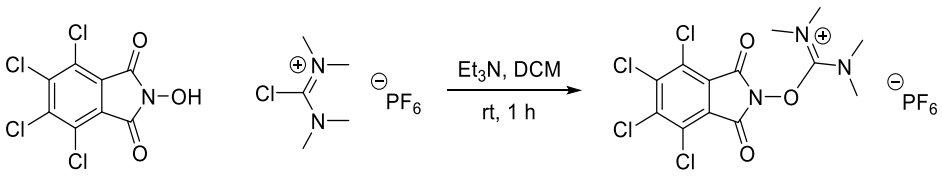


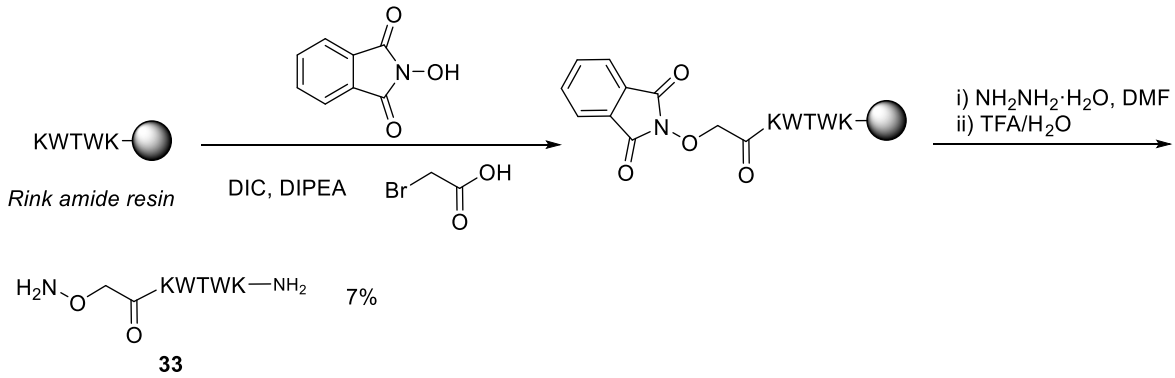
Figure 142. Oxyma adduct generated upon coupling of bromoacetic acid.

To overcome this undesired reactivity, a different coupling strategy was sought. The coupling agent CITU[332] is synthesised through the reaction between an *N*-hydroxyphthalimide centre and tetramethyluronium chloride to generate an uronium centre (Scheme 42). Its resemblance to a DIC-activated coupling inspired the idea to use *N*-hydroxyphthalimide as a substitute for Oxyma Pure in order to achieve the coupling of the phthalimide and the acid in one step.



Scheme 42. Synthesis of CITU.[332]

To achieve this, the resin-bound peptide was treated with 9.5 equiv. *N*-hydroxyphthalimide, bromoacetic acid (5 equiv.), DIC (4.5 equiv.) and DIPEA (6 equiv.) in DMF for 3 h at room temperature (Scheme 43). Analysis *via* a cleavage test and LCMS showed complete conversion to the product. The synthesis was continued with the deprotection of the *N*-hydroxyphthalimide with hydrazine hydrate (80% in DMF) to generate the *N*-alkoxy amine moiety, followed by cleavage from the resin using TFA and water (95/5%). This afforded **33** in a 7% overall yield after purification.



Scheme 43. Synthesis of the *C*-terminal fragment, 33.

After storing a sample of **33** in a plastic tube, it was observed that an adduct with a mass 40 Da over the product was generated. This was identified as an acetone-like adduct on the oxime (Figure 143) by mass spectroscopy. It has previously been observed that *N*-alkoxy amines are very reactive towards ambient traces of solvents like acetone and formaldehyde, as well as softeners in plastic tubes.[319,333] In fact, Duflocq *et al*.[319] have recently reported the use of propargyloxycarbonyl-protected *N*-alkoxy amine moieties for oxime ligation, which requires a Pd(II)-catalysed deprotection step before ligation but prevents the formation of undesired oxime bonds. In this case, it was found that if glassware was used for all manipulation of the peptide after cleavage and if the presence of acetone was avoided, only minor fractions of the undesired oxime were observed.

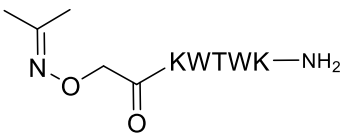
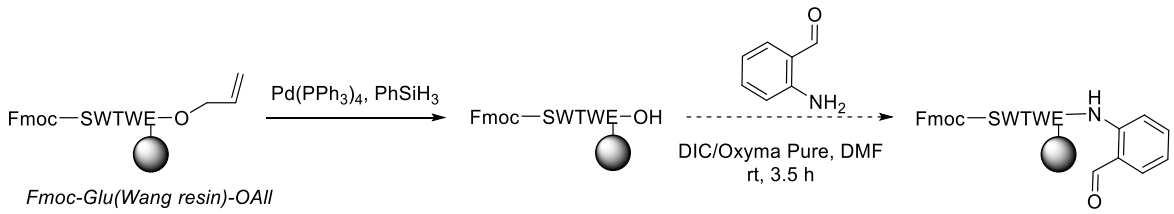


Figure 143. Proposed undesired adduct generated upon storage of 33 in plastic containers.

5.4.2 Synthesis of the *N*-terminal fragment 32

The synthesis of **32** was started on 0.1 mmol of Fmoc-Glu(Wang)-OAll resin (0.44 mmol/g loading). Standard microwave-assisted SPPS with DIC/Oxyma Pure activation was used to couple the native amino acids, with the *N*-terminus being left Fmoc-protected (Scheme 44). After deprotection of the allyl group using phenyl silane and palladium (0) tetrakis(triphenylphosphine), the coupling of 2-aminobenzaldehyde (1 equiv.) was attempted using DIC (1.2 equiv.) and Oxyma Pure (1.2 equiv.) in DMF, at room temperature for 3.5 h (Table 16, entry 1). A test cleavage was then performed, but this only showed the presence of uncoupled peptide.



Scheme 44. Synthetic route towards the *N*-terminal fragment, 32.

It was proposed that the low nucleophilicity of the amine in 2-aminobenzaldehyde could be the cause for no product being observed. Coupling conditions using HATU in DMF were also tested (Table 16, entry 6), as this coupling agent generates a more reactive intermediate which could favour the nucleophilic attack, but only an unidentified by-product was generated. Other coupling conditions were tested which are presented in Table 16.

Table 16. Conditions tested for the coupling of 2-aminobenzaldehyde.

Entry	Coupling conditions	Solvent	Temperature	Time	Equiv. of 2-amino benzaldehyde	Product after test cleavage
1	DIC/Oxyma 1.2 equiv.	DMF	rt	3.5 h	1	Starting material
2	DIC/Oxyma 10 equiv.	DMF	rt	24 h	5	Starting material
3	DIC/Oxyma 1.2 equiv.	DMF	50 °C	3.5 h	1	Starting material
4	DIC/Oxyma 10 equiv.	DCM	rt	24 h	5	Unknown by-products

Table 16. Conditions tested for the coupling of 2-aminobenzaldehyde.

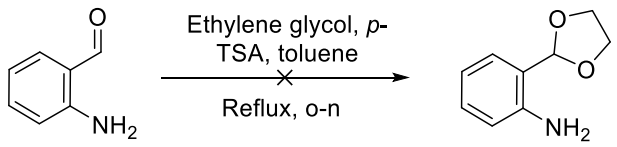
Entry	Coupling conditions	Solvent	Temperature	Time	Equiv. of 2-amino benzaldehyde	Product after test cleavage
5	Microwave-assisted coupling on synthesiser, DIC/Oxyma 5 equiv.	DMF	90 °C	3 min	5	Starting material
6	HATU 1.2 equiv., DIPEA 2.5 equiv.	DMF	rt	3.5 h	1	Starting material and unknown by-products
7	BOP 1.2 equiv., Et_3N[334]	DMF	rt	o/n	1	Unknown by-products
8	EDCI 1.1 equiv., DMAP 10%[335]	DCM	rt	24 h	1	Starting material
9	EEDQ 10 equiv.	DMF	rt	24 h	5	Starting material
10	EEDQ 10 equiv.	DCM	rt	24 h	5	Starting material

None of the tested conditions furnished the coupled product. This could be due to the low nucleophilicity of the amine or side reactions taking place with the aldehyde functionality. Protection of the aldehyde as an acetal was to be tested next, which would prevent potential side reactions from taking place during the coupling.

5.4.2.1 Acetal protection of the aldehyde

The initial strategy explored for the protection involved the use of ethylene glycol (dry, 1.5 equiv.) and catalytic *p*-toluenesulfonic acid (*p*-TSA) (0.1 equiv.) in toluene under a nitrogen atmosphere, to furnish the acetal-protected **34** (Scheme 45).[336,337] After reflux overnight no

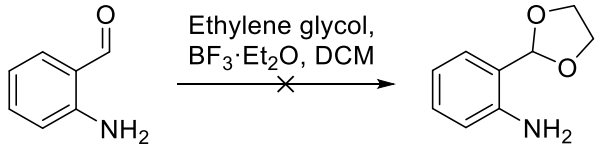
product mass was observed, and instead high-resolution mass spectrometry analysis revealed the presence of a mass of 371 Da, which corresponded to more than double of the mass of the desired product (165 Da). These conditions were tested again, this time with 0.2 equiv. of *p*-TSA and 5 equiv. of ethylene glycol, but the same undesired product was observed. It was suggested that *p*-TSA could be protonating the amine and causing side reactions, or that the formation of adducts prevented the formation of the desired acetal product.



34

Scheme 45. Initial strategy for the protection of the aldehyde.

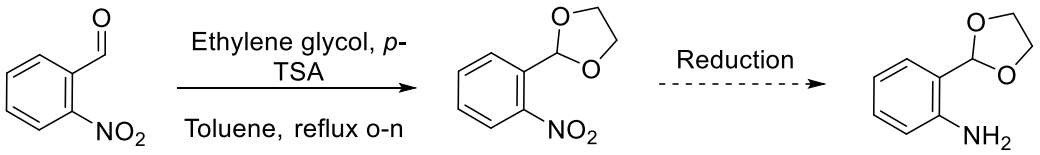
Different conditions were then tested, using boron trifluoride etherate (1.5 equiv.) and 4 equiv. of ethylene glycol in DCM from 0 °C to room temperature overnight (Scheme 46),[338] but only starting material was recovered.



34

Scheme 46. Second strategy for the aldehyde protection.

In order to prevent the protonation of the amine, the acetal protection was then tried on 2-nitrobenzaldehyde (Scheme 47). Using 0.2 equiv. of *p*-TSA and 10 equiv. of ethylene glycol in toluene afforded a quantitative yield of the desired product **35**.



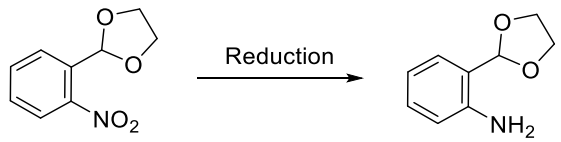
35 **34**

Scheme 47. Synthesis of the acetal-protected 35, followed by its reduction to 34.

5.4.2.2 Reduction of 35

Acid-free reducing conditions were then sought for the synthesis of **34** – the conditions tested are presented in Table 17. Using Pd/C and H_2 (entry 1) did not generate any product, and starting material was left unreacted. NMR analysis showed that a by-product with the deprotected aldehyde was generated (12% conversion), although its structure could not be identified. The selective reduction of nitro groups in presence of a carbonyl is often difficult – therefore indium and iron powder were also tested (entries 4 and 5 respectively). These conditions have been shown to be efficient in this context with good functional group tolerance,[339] but in this case the aldehyde was deprotected and 2-aminobenzaldehyde was recovered instead (100% conversion for entry 5 and 18% for entry 4, where 82% of the conversion afforded another aldehyde-containing by-product).

Table 17. Conditions tested for the reduction of 35.



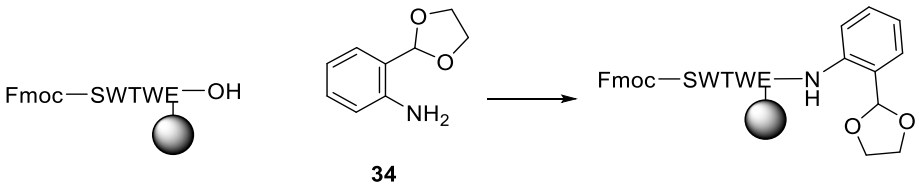
35 **34**

Entry	Reduction conditions	Base	Solvent	NMR conversion to 34
1	H_2, 10% Pd/C[337]	Et_3N	*Iso*-propyl alcohol (IPA)	0%
2	$HSiCl_3$[340]	DIPEA	MeCN	0%
3	Na_2S[341]	Et_3N	EtOH	Quantitative
4	In[339]	NH_4Cl	EtOH	0%
5	Fe[342]	NH_4Cl	EtOH	0%

The most suitable conditions found used sodium sulfide, which afforded quantitative conversion to **34** by NMR. When it was attempted to scale up the reaction, the crude purity worsened, therefore the reaction was performed under a nitrogen atmosphere. This afforded **34** in a 65% yield after purification.

5.4.2.3 Coupling of 34 onto the resin-bound peptide

With acetal-protected **34** in hand, conditions for its coupling to the resin-anchored peptide were investigated (Scheme 48, Table 18). A test cleavage was performed after coupling using 95% TFA, 5% H_2O to evaluate the success of the reaction.



34

Scheme 48. Coupling of 34 onto the resin-linked peptide.

Conditions using DIC/Oxyma Pure in DMF or DCM with heating did not show product formation, and product traces were observed when performing the coupling at room temperature for 24 h (Table 18, entry 2). Other coupling agents and the use of a microwave synthesiser did not improve the conversion. Other resin-compatible non-polar solvents were to be tested next in order to prevent hydrogen bond formation between the amine and the solvent – it was hoped that this would favour the nucleophilic attack of the amine. Traces of product were observed when using DIC/Oxyma Pure under reflux in dichloroethane and toluene (entries 7 and 8).

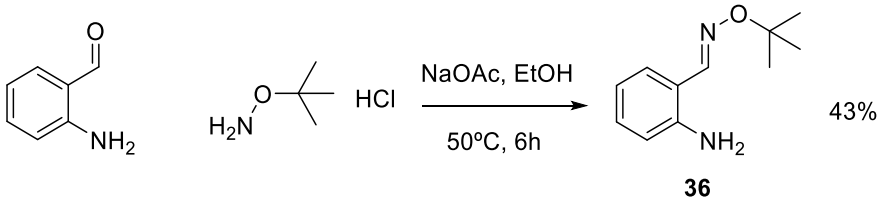
Table 18. Conditions explored for the coupling of 34.

Entry	Coupling conditions	Product after test cleavage
1	DIC/Oxyma 10 equiv., DMF, 50 °C 24 h	-
2	DIC/Oxyma 10 equiv., DCM, rt 24 h	Trace
3	DIC/Oxyma 10 equiv., DCM, 40 °C 24 h	-
4	Microwave-assisted coupling on CEM synthesiser, DIC/Oxyma 5 equiv., DMF	-
5	HATU 10 equiv., DIPEA, DMF, rt 24 h	-
6	EEDQ 10 equiv., DMF, rt 24 h	-
7	DIC/Oxyma 10 equiv., DCE, 80 °C 24 h	Trace
8	DIC/Oxyma 10 equiv., toluene, 90 °C 24 h	Trace

No sufficiently high-yielding conditions were found using this strategy, therefore a new protection strategy for the aldehyde was sought.

5.4.2.4 Oxime protection of the aldehyde

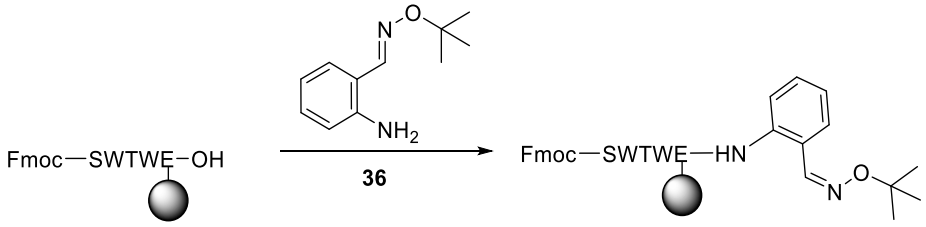
A different strategy for the protection of the aldehyde was investigated, namely the use of an oxime bond. **36** was obtained in a 43% yield through the reaction between 2-aminobenzaldehyde and O-^{t}Bu-hydroxylamine hydrochloride (2 equiv.) in EtOH, at 50 °C for 6 h (Scheme 49).[343]



Scheme 49. Synthesis of 36.

5.4.2.5 Coupling of 36 onto resin-bound peptide

The coupling of **36** was attempted using various conditions (Scheme 50, Table 19). DIC/Oxyma Pure activation in DMF at room temperature (entry 1) showed consumption of the carboxylic acid-containing peptide starting material, but no coupled product. The use of DCM as the solvent and reflux conditions showed no improvement.



Scheme 50. Coupling of 36 on the resin-anchored peptide.

Other coupling agents, namely HATU and EEDQ were also tested (entries 3 and 4 in Table 19), but no product formation was observed. The carboxylic acid-containing peptide starting material was not consumed either, and no coupling agent adducts were observed.

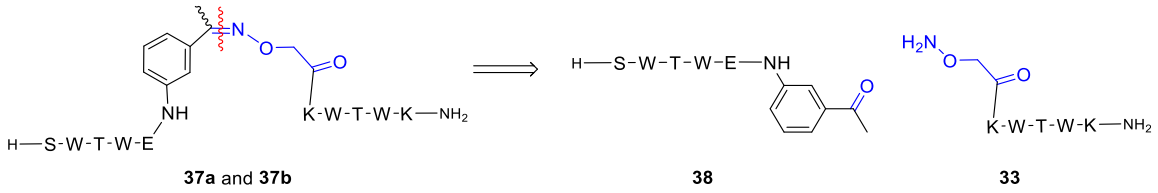
Table 19. Conditions tested for the coupling of 36. All reactions were performed for 24 h.

Entry	Coupling conditions	Solvent	Temperature	Product after test cleavage
1	DIC/Oxyma 10 equiv.	DMF	rt	-
2	DIC/Oxyma 10 equiv.	DCM	40 °C	-
3	HATU 10 equiv., DIPEA	DMF	rt	-
4	EEDQ 10 equiv.	DMF	rt	-

The very low yields of these coupling reactions were attributed to the low nucleophilicity of the amine in the small molecule building block. Therefore a different substitution of the aromatic ring was to be explored next.

5.5 *Meta*-substituted oxime mimic

The second target pursued consisted of a *meta*-substituted aromatic ring instead of the *ortho*-substituted ring, which would hopefully improve the nucleophilicity of the amine while still providing a good geometry for the generation of the turn. Due to 3'-aminobenzaldehyde not being available, it was decided that 3'-aminoacetophenone would be used instead to generate the ketoxime-containing conjugate (Scheme 51).



Scheme 51. Oxime-linked target and the two peptide substrates for the ligation.

Molecular dynamics simulations were run on both stereoisomers by Dr Drew Thomson, which showed that the *Z*-oxime maintained the β-hairpin structure (Figure 144). Conversely, the *E*-oxime appeared to break the TrpZip structure, giving rise to an almost completely unfolded system.

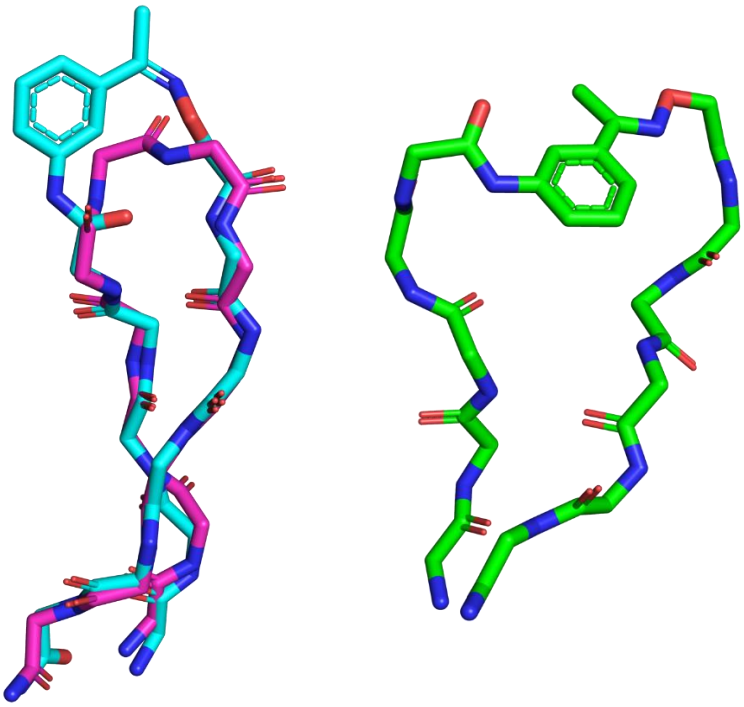


Figure 144. Left: Overlay of a frame extracted from the MD simulation of the *Z*-oxime (in blue) and of the NMR structure of TrpZip1 (in magenta, PDB 1LE0). Right: frame extracted from the MD simulation of the *E*-oxime.

The potential difference in folding between the stereoisomers was further confirmed by the distances between the oxygen in the C=O(*i*) and the proton in NH(*i*+*3*) that were extracted from the MD trajectories. The average distance along the simulation for the *Z*-oxime was of 2.0 Å, whereas for the *E*-oxime it was of 7.2 Å (Figure 145). This further suggested that the system containing the *Z*-oxime β-turn mimic would adopt a β-hairpin conformation with the required hydrogen bond, whereas the conjugate with the *E*-oxime mimic would not adopt a folded conformation.

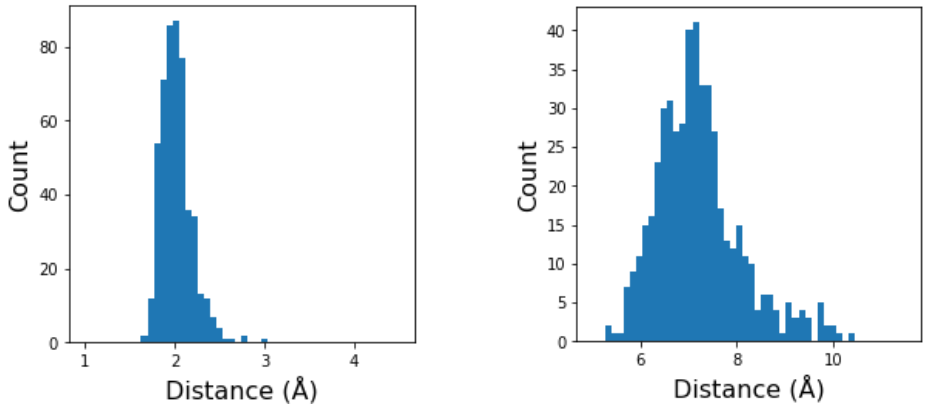
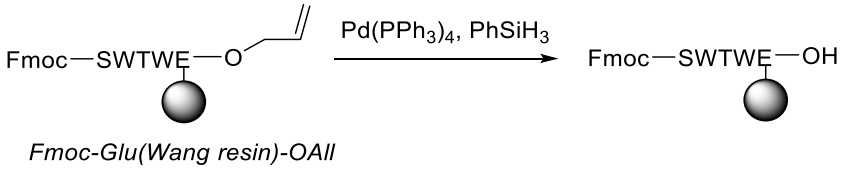


Figure 145. Histograms of the distances between the oxygen in the C=O(*i*) and the proton in NH(*i*+*3*) for the *Z*-oxime (left) and the *E*-oxime (right).

5.5.1 Synthesis of the peptide fragments 33 and 38

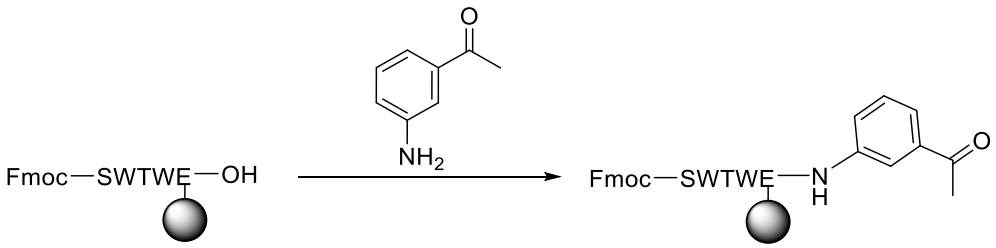
The synthesis of the *C*-terminal fragment **33** was performed as described in Section 5.4.1. The synthesis of the *N*-terminal fragment, **38**, was performed on Fmoc-Glu(Wang)-OAll resin (0.44 mmol/g loading). The synthesis was started as described in Section 5.4.2, using microwave-assisted SPPS followed by removal of the allyl protecting group (Scheme 52).



Scheme 52. Synthetic route towards the *N*-terminal fragment 38.

The coupling of 3'-aminoacetophenone was then attempted (Table 20). The first conditions tested employed DIC/Oxyma Pure in both DMF and DCM, using a high number of equivalents as well as an extended reaction time, at room temperature (entries 1 and 2 in Table 20). This afforded some conversion to the desired product, although traces of the non-coupled starting material were still present. Conditions using HATU activation were then tested, but no product formation was observed.

Table 20. Coupling conditions tested for the coupling of 3'-aminoacetophenone on the resin-bound peptide.



Entry	Coupling conditions	Product after test cleavage
1	DIC/Oxyma 10 equiv., DMF, rt 24 h	50% conversion
2	DIC/Oxyma 10 equiv., DCM, rt 24 h	75% conversion
3	HATU, DIPEA 10 equiv., DMF, rt 24 h	No product formation
4	DIC/Oxyma 10 equiv., DCE, 80 °C 24 h	95% conversion

More forcing conditions were then sought: analogous to the coupling of **34** described in section 5.4.2.3, it was reasoned that a non-polar solvent would favour the nucleophilic attack of the amine. The reaction was performed under reflux in dichloroethane (Table 20, entry 4), which provided 95% conversion to the desired product. Because these conditions involved a high temperature and extended reaction time, an additional product, possibly due to racemisation of the *C*-terminal amino acid was observed (Figure 146), but this was a minor product and could be separated from the desired peptide.

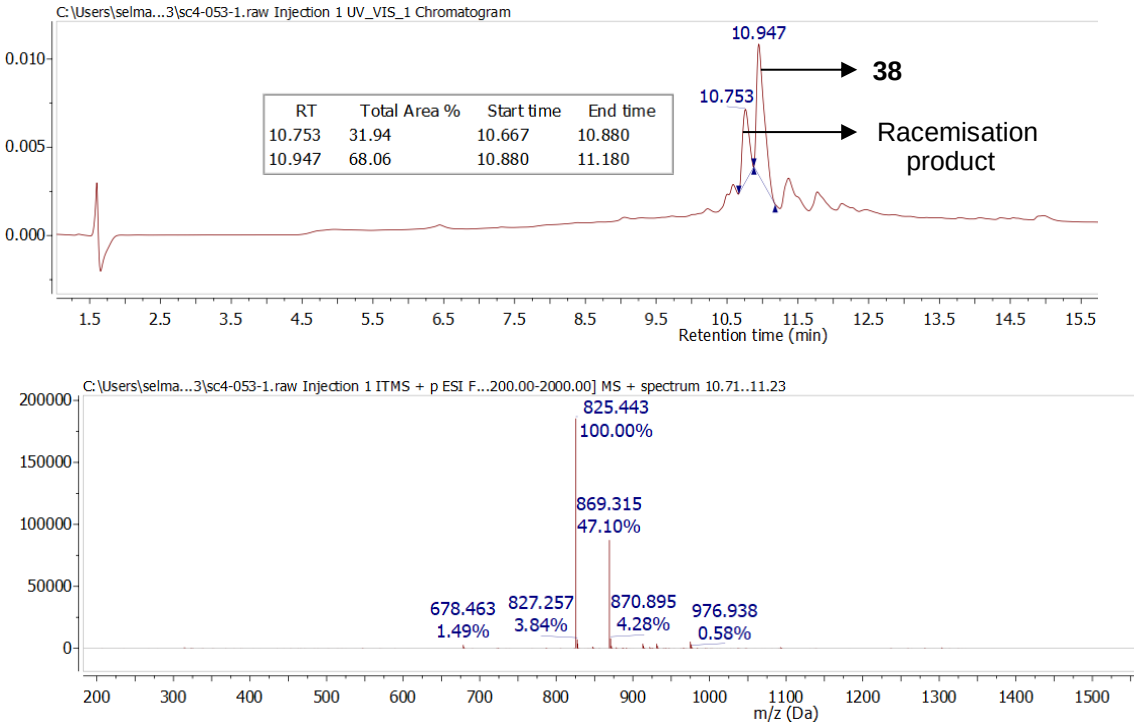
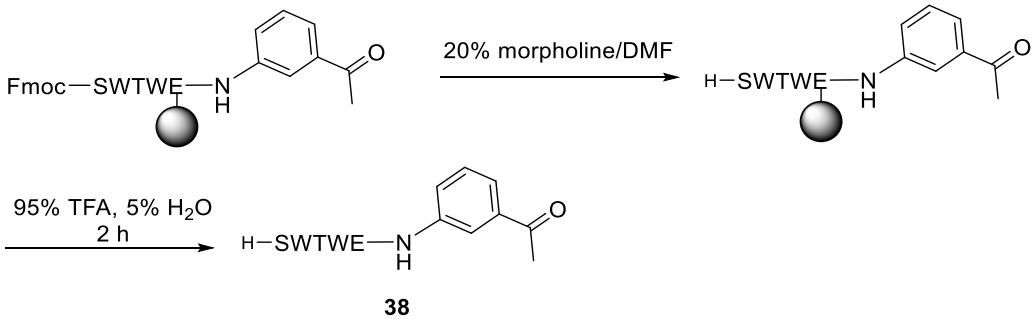


Figure 146. LCMS trace of the test cleavage after coupling of 3'-aminoacetophenone and Fmoc deprotection. The top trace is the UV-Vis chromatogram, and the bottom corresponds to the mass spectrum. The peak at 10.75 min was identified as the racemisation product whilst the peak at 10.95 min corresponds to 38.

The synthesis was then continued by the Fmoc deprotection of the *N*-terminus and cleavage under standard conditions (Scheme 53), which afforded **38** in a 6% overall yield.

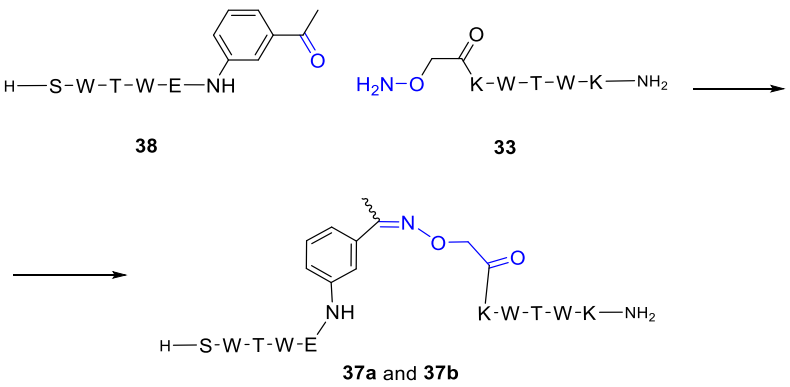


38

Scheme 53. Final steps in the synthesis of 38.

5.5.2 Oxime ligation towards 37a and 37b

The ligation between the peptide partners to generate **37a** and **37b (**Scheme 54**)** was attempted at a concentration of 1 mM under four different conditions. Firstly, conditions in aqueous buffer were explored: a NH_4OAc buffer was used (Table 21, entry 1), which provided an acidic pH that should favour the reaction. The solubility of the peptides was limited, thus 10% of DMF was added. After monitoring the reaction *via* HPLC for a week, 25% of product was observed. The low conversion could be due to the slow reaction rate of ketoxime formation, but also because of the limited solubility of the peptides in the buffer.



Scheme 54. Ligation between 38 and 33 to produce 37a and 37b.

Nucleophilic catalysis was then employed to try to improve the reaction rates (Table 21). Two strategies were tested, one using 10 equiv. of aniline and another using *p*-phenylenediamine (*p*-PDA), both in NH_4OAc buffer. When using aniline (entry 2), 17% of product was generated, not providing an improvement in comparison to the non-catalysed conditions. When *p*-PDA was used (entry 3) 12% of product was generated, but a significant amount of a by-product (58%) was formed. This by-product had a molecular weight of 859, which could correspond to an acetone-like adduct on the *N*-alkoxy amine-containing starting material (Figure 147). The purity of the *N*-alkoxy amine-containing peptide was confirmed and no adduct was present before the ligation reaction. The *p*-PDA used was analysed as well, but no acetone traces were observed. This adduct was also observed in the other ligation tests but only traces were present.

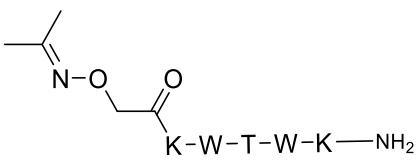


Figure 147. Proposed acetone-like adduct on 33.

New ligation conditions were thus sought, turning to MeOH/AcOH 1:1 as the solvent system (Table 21, entry 4). This solvent greatly improved the solubility of the starting materials, leading to good conversion to the desired products in 3 days, with 70% of **37a** and **37b** being generated in total by HPLC. Therefore these were the chosen conditions for this ligation. **37a** and **37b** have been named as such according to the order in which they are eluted from the column during RP-HPLC purification.

Table 21. Conditions tested for the ligation to generate 37a and 37b.

Entry	Substrate concentration	Solvent	Equiv. of catalyst	Product generated in HPLC trace
1	1 mM	NH_4OAc buffer 0.1 M pH 4.5, 10% DMF	-	25%
2	1 mM	NH_4OAc buffer 0.1 M pH 4.5, 10% DMF	10 equiv. aniline[195]	17%
3	1 mM	NH_4OAc buffer 0.1 M, pH 4	10 equiv. *p*-PDA[323]	12%
4	1 mM	MeOH/AcOH 1:1	-	70%

HPLC monitoring of the ligation using MeOH/AcOH as the solvent system showed complete consumption of **33** after 3 days (Figure 148). After that time, only more **37b** was generated, which could be due to the isomerisation of **37a**. The proportion between the stereoisomers after 6 days was of 23% **37a** and 77% **37b**.

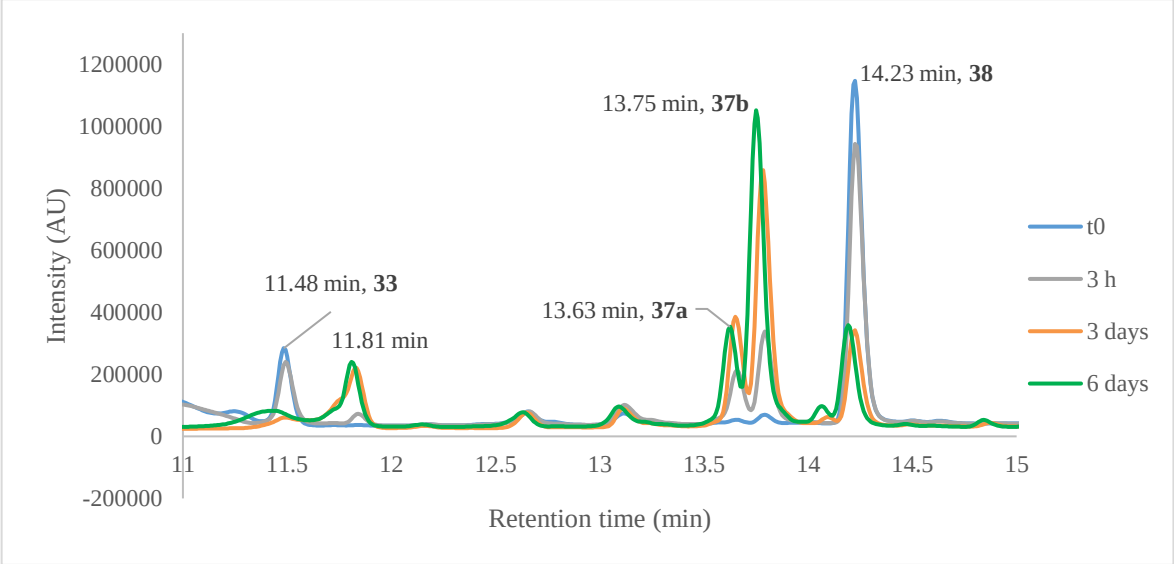


Figure 148. HPLC monitoring of the ligation towards 37a and 37b.

5.5.3 Isomer interconversion tests in TFA

During purification of the ligated product it was observed that although the isolated fractions showed one isomer peak on the RP-HPLC, HPLC analysis once lyophilised showed that a fraction of the other stereoisomer was present. This could be caused by the 0.1% TFA present in the solvent system, which could cause the isomerisation of the oxime bond. Multiple mechanisms have been identified as possible routes for the acid-catalysed isomerisation of imine derivatives. One proposed pathway proceeds *via* hydrolysis followed by recondensation, while another suggests the rotation of the C=N bond after protonation of the nitrogen to the iminium form.[344,345] The appearance of the other stereoisomer during purification was only observed to take place with the first eluting stereoisomer, **37a**, but not with the second, **37b**. To confirm that isomerisation was taking place, a sample of each stereoisomer was re-dissolved in H_2O/MeCN 1:1 with 0.1% TFA and monitored through HPLC.

In the case of the first eluting isomer, **37a**, 50% of interconversion to **37b** was observed after 2 days, with almost complete conversion after 8 days (Figure 149). An unknown by-product was also generated, with a retention time of 14.3 min.

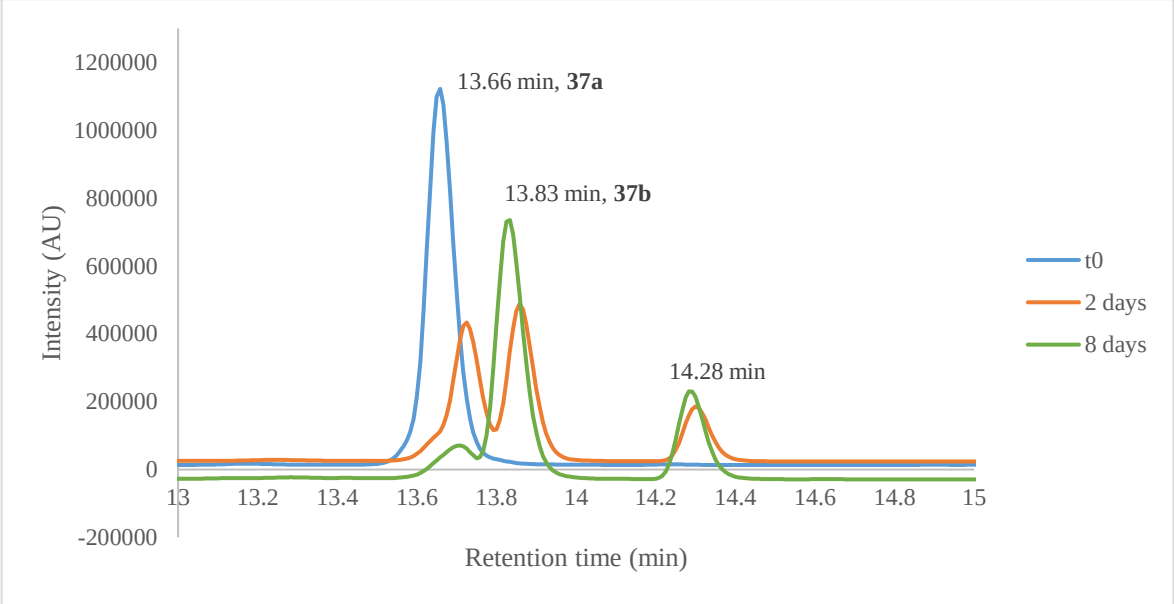


Figure 149. Interconversion test of 37a.

When the same test was performed with **37b**, no interconversion took place (Figure 150). This indicated that the stereoisomers were under thermodynamic control (oxime isomers are known to equilibrate under acidic conditions[317,345]) and that **37b** is the thermodynamically stable product, which could point towards **37a** corresponding to the *Z*-oxime stereoisomer and **37b** to the *E*-oxime.

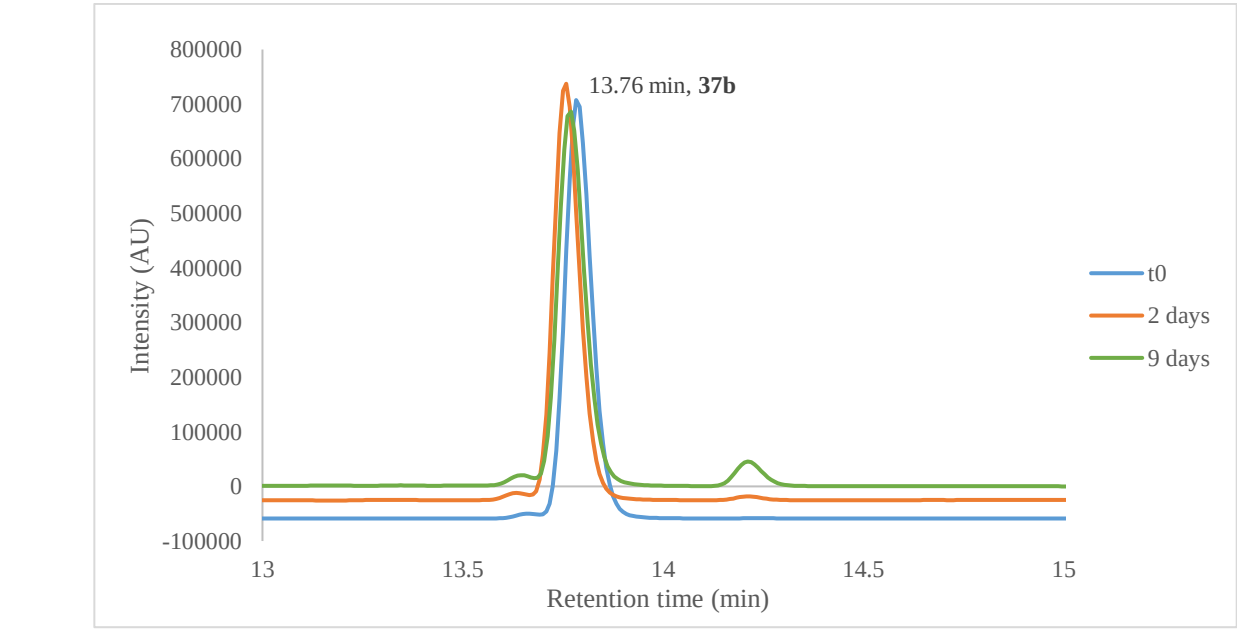


Figure 150. Interconversion test of 37b.

5.5.4 Circular dichroism analysis of 37a and 37b

In order to evaluate the secondary structure and thermal stability of the conjugates, circular dichroism experiments were performed. These experiments were run in the far UV using analogous conditions to those used in Chapter 3: in sodium phosphate buffer (20 mM) at pH 7, with a peptide concentration of 0.05 mg/mL (30 μM). The folding of the conjugates was compared to that of the **TrpZip1** control. It was attempted to run experiments in the near UV at a 1 mg/mL peptide concentration, but the solubility of the peptides was very limited and precipitation was observed. The peptides were also dissolved in HEPES buffer (10 mM, pH 7), expecting the lower ionic strength to aid solubility, but the solubility was still low. Therefore only experiments in the far UV region were used to evaluate the folding and thermal stability of these constructs.

The far UV CD showed that both stereoisomers adopted β-hairpin structures, with the characteristic negative band at 212 nm being present as well as the 228 nm positive band indicative of tryptophan stacking (Figure 151).[141,276] Interestingly, although **37b** was the thermodynamically stable product it was considerably less folded than **37a**. In order to confirm that this difference did not arise from a variation in concentration or precipitation of the peptide, the CD experiments were repeated in HEPES buffer and the concentration was measured before and after the experiment through UV-Vis absorbance. This provided the same folding curves as in the sodium phosphate buffer. In addition, no rise in high tension (HT) was observed at low wavelengths, which would be indicative of scattering – this further confirmed that no precipitation was taking place.

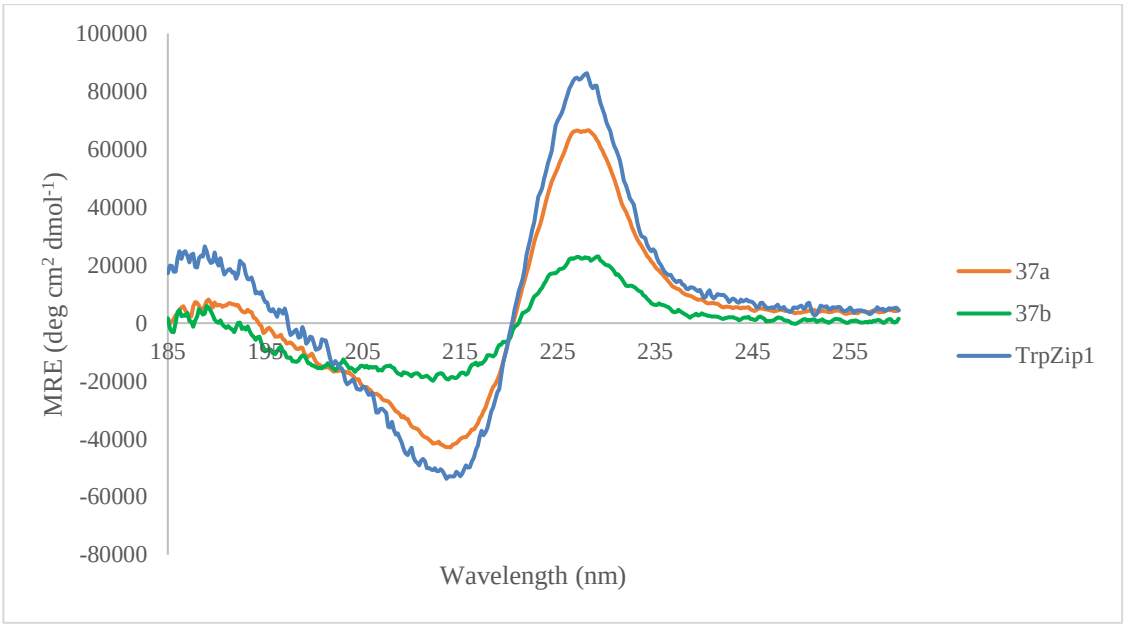


Figure 151. Far UV folding of 37a and 37b, in comparison to TrpZip1.

The thermal stability of the conjugates was evaluated by monitoring the change in CD upon gradual heating. For this, the MRE was monitored at 228 nm every 1 °C, from 5 to 80 °C (Figure 152). The folding curves before and after heating were superimposable for both constructs, showing that the unfolding was a reversible process. The denaturation curves for **37a** and **37b** had lower MRE values than that of **TrpZip1**, and were shallower in shape. In particular, the melting curve of **37b** had low MRE values and resembled the end of a sigmoid, confirming the low degree of folding of this peptide even at low temperatures. The shape of the curve of **37b** also indicated that the transition point was at lower temperatures than 5 °C, and therefore a melting temperature for this sequence could not be measured. This further emphasised the low stability of this construct, and suggested that **37b** could correspond to the *E*-oxime by comparing to the MD simulations.

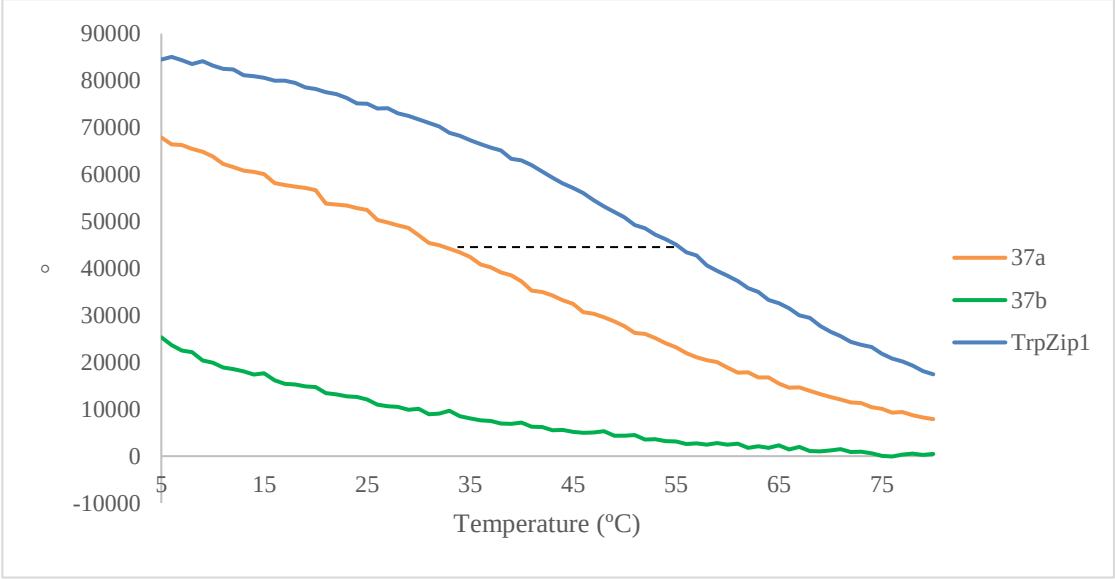


Figure 152. Thermal denaturation experiments for all peptides in study. The black dashed line indicates an estimation of the midpoint of the curve for TrpZip1 and the point at which 37a adopts the same degree of folding.

The thermal denaturation curves of **37a** and **TrpZip1** were different in MRE values, with **37a** being more unfolded at the same temperature. To illustrate this, if the midpoint of the curve for **TrpZip1** is estimated (by finding the halfway point between the maximum MRE value and zero, black dashed line in Figure 152), at 57 °C, **37a** would adopt the same degree of folding at 34 °C, a temperature 23 °C lower.

The thermal denaturation curves of **37a** and **TrpZip1** were not sigmoidal and a clear transition midpoint could not be observed from the first derivative of the curves (Figure 153), therefore the fraction of folding was used to estimate a melting temperature.

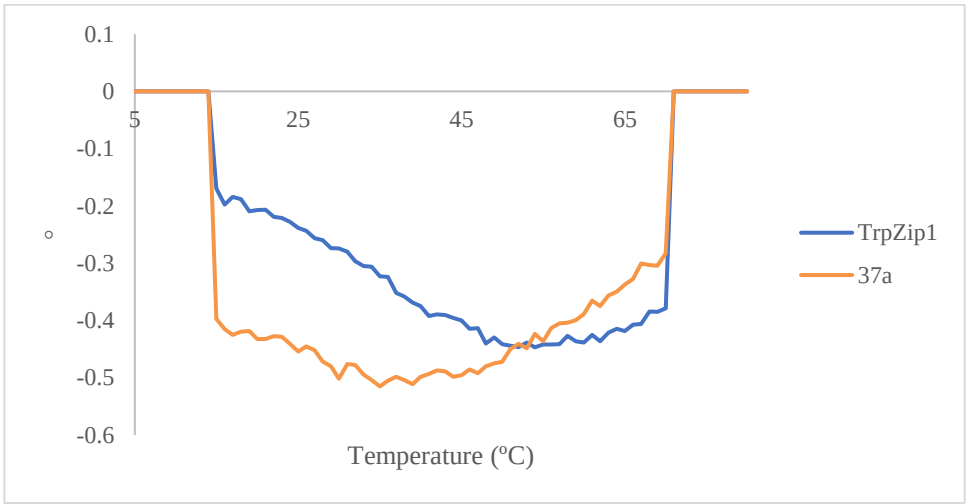


Figure 153. First derivatives of the melting curves for 37a and TrpZip1.

The fraction of folding was calculated[277] for the two peptides by adopting an MRE of zero as the unfolding process endpoint and the MRE value of **TrpZip1** at 5 °C as the fully folded form, as **37a** appeared to be more unfolded than the control at that temperature. This afforded the fraction of folding curves presented in Figure 154, which confirmed that **37a** was considerably less folded than the control sequence.

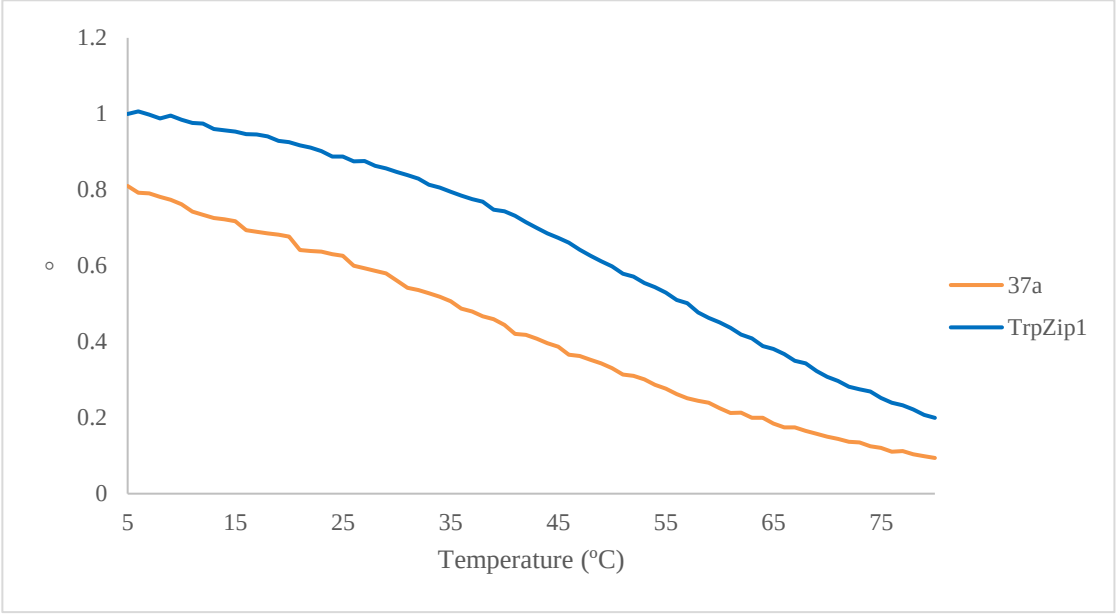


Figure 154. Fraction of folding for all peptides in study.

The T_m estimated from the fraction of folding curves indicated that the stability of **37a** was lower than that of **TrpZip1** by 22 °C (Table 22).

Table 22. T_m calculated for the peptides under study.

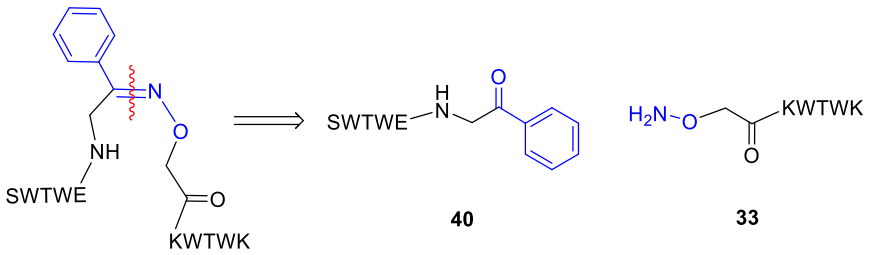
Sequence	T_m (°C)
TrpZip1	57
37a	35

The identity of the two stereoisomers was inferred by comparing the CD data with the information obtained from the MD simulations. The simulations predicted that the *E*-oxime would induce an almost unfolded conformation, a behaviour observed by CD analysis for **37b**. On the other hand, **37a** behaved as predicted for the *Z*-oxime, adopting a folded β-hairpin conformation.

Overall, this design afforded two stereoisomers with completely different behaviours: one with a *Z*-oxime which successfully mimics the β-turn and maintains the β-hairpin structure albeit with a lower stability than the control, and another with an *E*-oxime that adopts a largely unfolded β-hairpin.

5.6 Phenyl-functionalised oxime mimic

A different target was pursued, which was based on a ketoxime with a phenyl substituent. This system was chosen because it would provide less preorganisation in comparison to the aromatic ring-containing analogues, therefore it would be interesting to compare the turn-inducing properties of both systems. Two stereoisomers would be generated upon oxime ligation, of which the *E*-oxime would favour the turn geometry (Scheme 55). In addition, this system would provide a versatile backbone onto which more functionalisation could be added.



Scheme 55. Ligated target, 39, and the two peptide substrates for the ligation.

Energy minimisation experiments were performed for the two stereoisomers in the system, which showed maintenance of the β-hairpin structure (Figure 155). The distance between the oxygen in the carbonyl of the residue preceding the mimic and the following nitrogen was of 4.0 Å for both stereoisomers, suggesting a slight distortion in the turn.

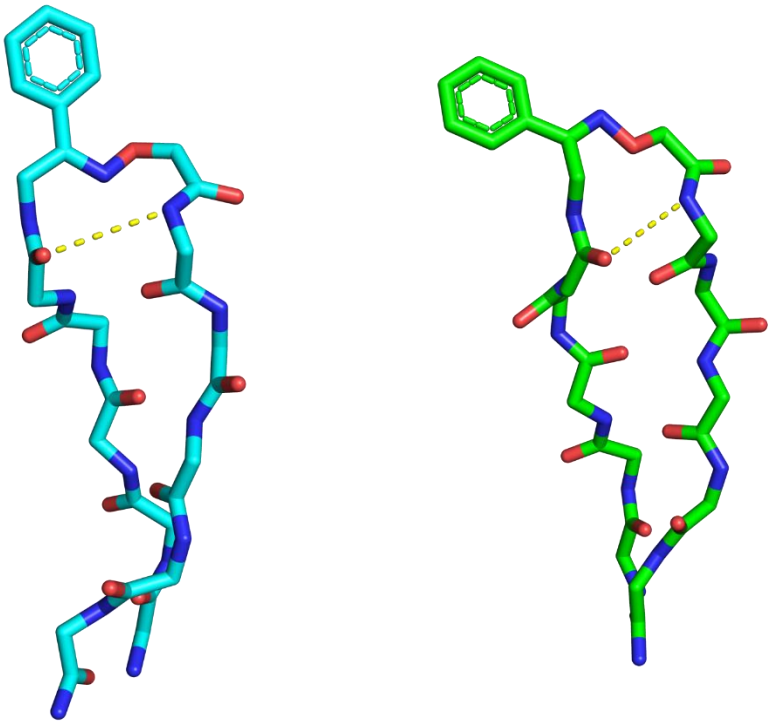
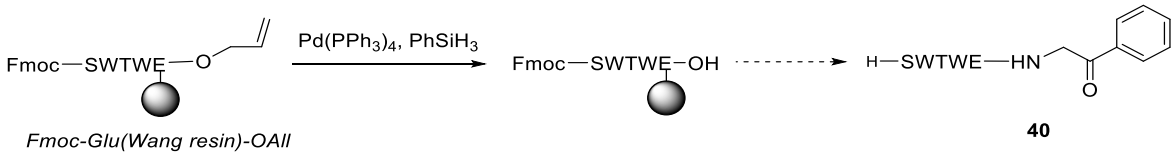


Figure 155. Energy minimisation models obtained for *Z*-39 (left) and *E*-39 (right), showing only the backbone atoms and the *i* to *i+3* like hydrogen bond in yellow.

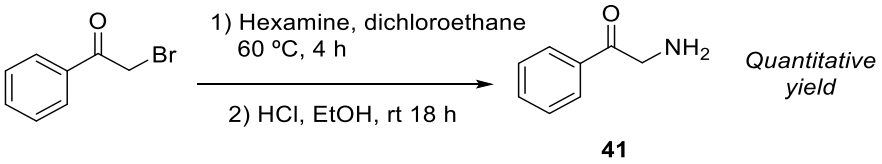
5.6.1 Synthesis of the peptide fragments 33 and 40

The synthesis of the *C*-terminal fragment **33** was performed as described in Section 5.4.1. The *N*-terminal fragment, **40**, was synthesised on Fmoc-Glu(Wang)-OAll resin (0.44 mmol/g loading). The synthesis was started as described in Section 5.4.2, using microwave-assisted SPPS to install the native amino acids followed by removal of the allyl protecting group (Scheme 56).



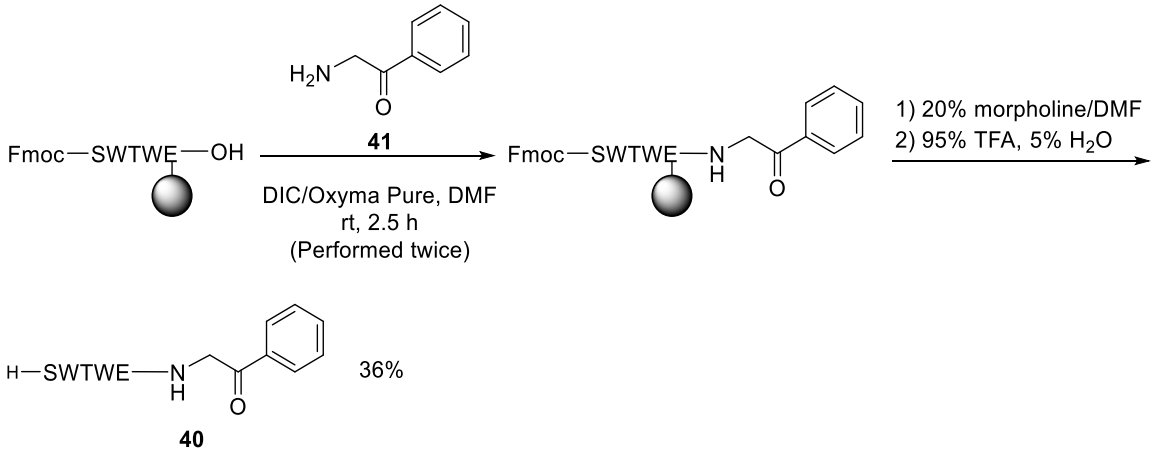
Scheme 56. First stage towards the synthesis of 40.

The ketone fragment required for coupling was synthesised through the Délepine reaction, following conditions described by Yang *et al*. (Scheme 57).[346] For this, hexamine was added onto 2-bromo-acetophenone under dry conditions and the reaction mixture was stirred at 60 °C for 4 h. After the acid hydrolysis of the salt generated, **41** was obtained in a quantitative yield without need for purification.



Scheme 57. Synthetic strategy towards 41.

Compound **41** was then coupled onto the resin-bound peptide using DIC/Oxyma Pure activation in DMF, at room temperature for 2.5 h (Scheme 58). However, LCMS analysis of a test cleavage sample showed that the starting material was still present in a 1:1 ratio to the desired product – therefore a double coupling was performed. This afforded complete conversion to the coupled product **40**.

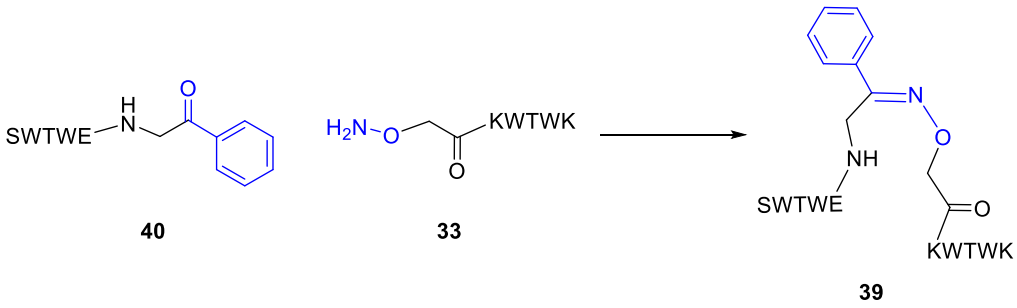


Scheme 58. Coupling of 41 onto the resin-linked peptide to generate 40.

After Fmoc deprotection and cleavage using 95% TFA and 5% H_2O, **40** was obtained in a 36% yield after purification.

5.6.2 Ligation step towards 39

The ligation between the two peptide partners **33** and **40** to furnish **39** was initially attempted under the conditions that were used for the other mimics, in MeOH/AcOH at a concentration of 0.75 mM (Scheme 59, entry 1 in Table 23). However, this did not result in product formation after two days. Aniline (10 equiv.) was then added in order to accelerate the reaction, but no product formation was observed. A screening of different conditions was then performed, which are presented in Table 23. All reactions were performed at room temperature and using a 1:1 ratio between the peptide fragments unless specified. DMF (5% of the volume of the solution) was added for all tests done in aqueous buffers, as solubility issues were observed for **40**.



Scheme 59. Ligation between 40 and 33 to generate 39.

A control ligation was performed in NH_4OAc buffer (entry 3) at a pH of 4.5 and without a nucleophilic catalyst, which only showed traces of product after 5 days of reaction. Catalysed conditions were then explored, using 10 equiv. of aniline (entry 4) and *p*-phenylenediamine (entry 5), but no improvement was observed. The poor catalytic ability of aniline in ketoxime ligations has previously been observed in the literature[319,324] – it was therefore decided to explore *p*-PDA further.

Table 23. Conditions tested for the ligation between 33 and 40 to furnish 39.

Entry	Ligation solvent	Buffer concentration	Catalyst	Peptide concentration	Product by LCMS analysis
1	MeOH/AcOH 1:1	-	-	0.75 mM	No product formation in 2 days
			After 2 days 10 equiv. aniline		No product formation after o-n
2	MeCN/H_2O 7:3, 1% TFA[309]	-	-	1 mM	Product traces after o-n
3	NH_4OAc buffer pH 4.5, 5% DMF	0.1 M	-	1 mM	Product traces after 5 days
4	NH_4OAc buffer pH 4.5, 5% DMF	0.1 M	10 equiv. aniline[195]	1 mM	Product traces after 5 days
5	NH_4OAc buffer pH 4, 5% DMF	0.1 M	10 equiv. *p*-PDA[323]	1 mM	Product traces after 4 days

Table 23. Conditions tested for the ligation between 33 and 40 to furnish 39.

6	NH_4OAc buffer pH 4, 5% DMF	0.1 M	10 equiv. *p*-PDA[323]	3.5 mM, 1.3 equiv. of **33** to 1 equiv. of **40**	40% conversion of **40** after 6 days
7	Potassium phosphate buffer pH 7.4, 5% DMF	20 mM	10 equiv. *p*-PDA[323]	1 mM	No product formation
8	Potassium phosphate buffer pH 4.5, 5% DMF	20 mM	200 equiv. $Sc(OTf)_3$ and *p*-PDA[347]	0.5 mM	Product traces after 6 days

It was attempted to ligate the peptide fragments in a ratio of 1.3:1 of **33** to **40**, as well as employing a higher concentration of 3.5 mM of each peptide (entry 6), which showed 40% conversion of **40** to **39** after 6 days. This highlighted how a higher concentration could favour the reaction, and that an excess of the *N*-alkoxy amine-containing fragment could be advantageous for these ligations. Formation of the previously mentioned acetone-like adduct was observed in all of the ligation tests performed (30-40% of the conversion of **33**), but it was a minor by-product in this test. It is possible that some of **33** was being consumed to generate the acetone-like adduct instead of remaining available for reaction, and that a higher concentration and more equivalents prevent this from taking place.

A study performed by Wendeler *et al*.[323] showed that *p*-substituted anilines have increased catalytic activity at pH 7 for oxime ligations, and thus a ligation was performed in a potassium phosphate buffer with 10 equiv. *p*-PDA (entry 7). However, this did not generate the desired product.

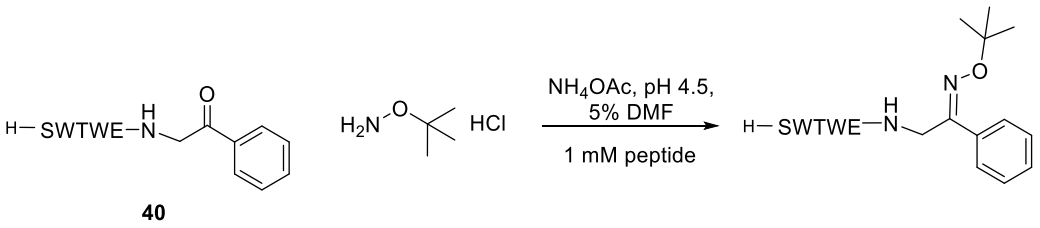
Cistrone *et al*. have shown that $Sc(OTf)_3$ used as a co-catalyst with *o*-PDA in oxime ligations increased the reaction rate up to one order of magnitude compared to *o*-PDA alone.[347] A test in potassium phosphate buffer was performed, using an excess of $Sc(OTf)_3$ and *p*-PDA (entry 8), but this did not provide an improvement.

Although one set of conditions was found that provided 40% conversion (entry 6), this was still a slow reaction. Ligation tests between the peptide partners and small molecules were

to be performed next, in order to confirm that the fragments could react to generate the oxime bond and to ascertain why the conjugation was slow.

5.6.3 Ligation tests with small molecules

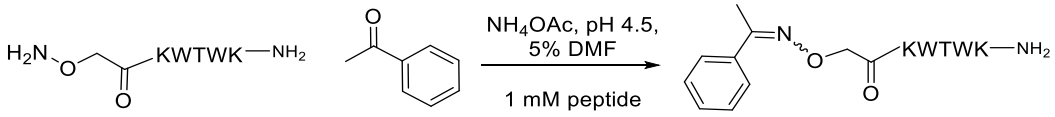
In order to study the behaviour of the two peptide partners, ligation tests were to be run between each peptide substrate and a small molecule equivalent to its partner. The first test performed was between **40** and O-^{t}Bu-hydroxylamine (8 equiv.), in a NH_4OAc buffer (0.1 M) at pH 4.5 without a catalyst (Scheme 60). HPLC monitoring of the reaction showed that after 5 days 40% of **40** was consumed to generate the ligated product.



40

Scheme 60. Ligation test between 40 and O-^{t}Bu-hydroxylamine.

The second test was between **33** and acetophenone (8 equiv.), which was also performed in a NH_4OAc buffer (0.1 M, pH 4.5) (Scheme 61). This showed complete product formation in the first 3 h. After this time, a small fraction of **33** was remaining, which was consumed to generate an unknown by-product.



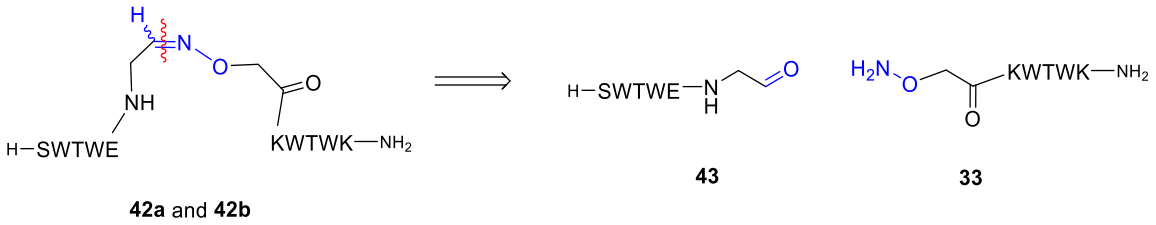
33

Scheme 61. Ligation test between 33 and acetophenone.

These ligation tests confirmed that the carbonyl-containing peptide fragment was causing the slow reactivity. This could be due to the carbonyl substrate being a ketone (as they usually have slower reaction rates in comparison to aldehydes[312]) but also due to it being bound to the peptide chain. Due to the explored ligation conditions not being sufficiently fast and the lack of a suitable catalyst, a new mimic structure was to be designed. This new target would contain an aldehyde moiety in order to favour reactivity.

5.7 Aliphatic oxime mimic

In order to increase the speed of the ligation reaction, the modified design would contain an aldehyde moiety on the *N*-terminal fragment (Scheme 62). This would remove the phenyl functionality present in **39**, which would afford a more flexible system with less preorganisation.



42a and **42b** **43** **33**

Scheme 62. Aliphatic oxime-containing target, 42a and 42b, and the two peptide substrates for the ligation.

Molecular dynamics simulations were run on both stereoisomers by Dr Drew Thomson, which showed that both conjugates maintained the β-hairpin structure in **TrpZip1** (Figure 156).

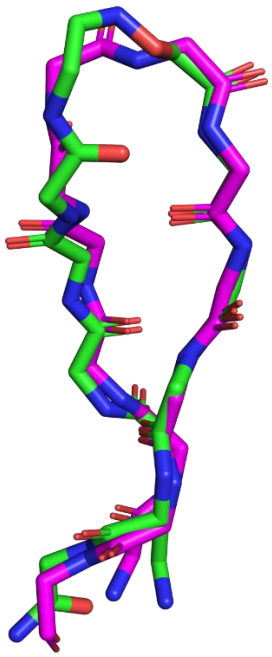


Figure 156. Overlay of a frame extracted from the MD simulation of the *Z*-oxime (in green) and of the NMR structure of TrpZip1 (in magenta, PDB 1LE0).

The distances between the oxygen in the C=O(*i*) and the proton in NH(*i*+3) were extracted from the trajectories, which were found to be a mean of 4.7 Å for the *Z*-oxime and a mean of 5.2 Å for the *E*-oxime (Figure 157). These were longer than the distances expected for a

hydrogen bond. By observing the simulations performed, it was found that the *Z*-oxime established and maintained the hydrogen bond towards the end of the trajectory. The *E*-oxime did not establish the hydrogen bond at any point and instead adopted a bulge-like geometry at the turn.

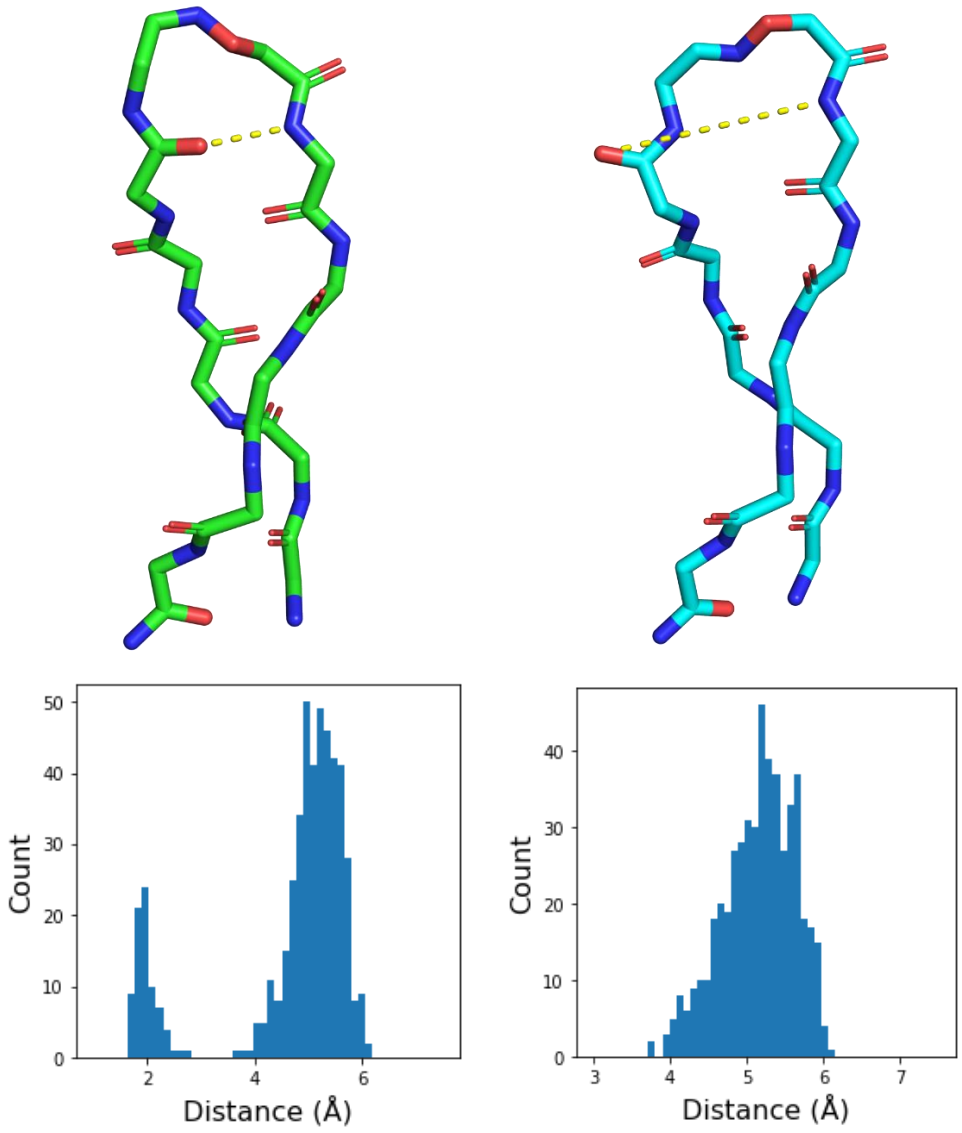


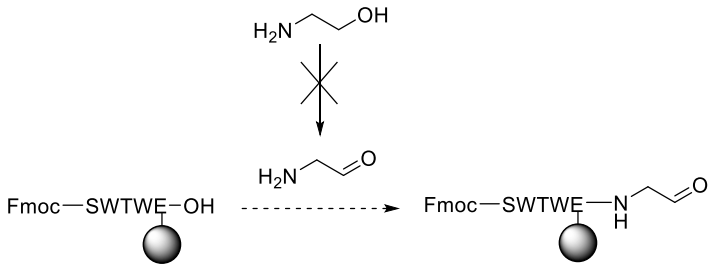
Figure 157. Top: frames extracted from the MD simulations of the *Z*-oxime (green) and the *E*-oxime (blue), showing the distance measured in yellow. Bottom: histograms of the distances between the oxygen in the C=O(*i*) and the proton in NH(*i+3*) for the *Z*-oxime (left) and the *E*-oxime (right).

5.7.1 Synthesis of the peptide fragments 33 and 43

The synthesis of the *C*-terminal fragment **33** was performed as described in Section 5.4.1. The synthesis of the *N*-terminal fragment was performed using Fmoc-Glu(Wang)-OAll resin (0.44 mmol/g loading). The synthesis was started as described in Section 5.4.2, using

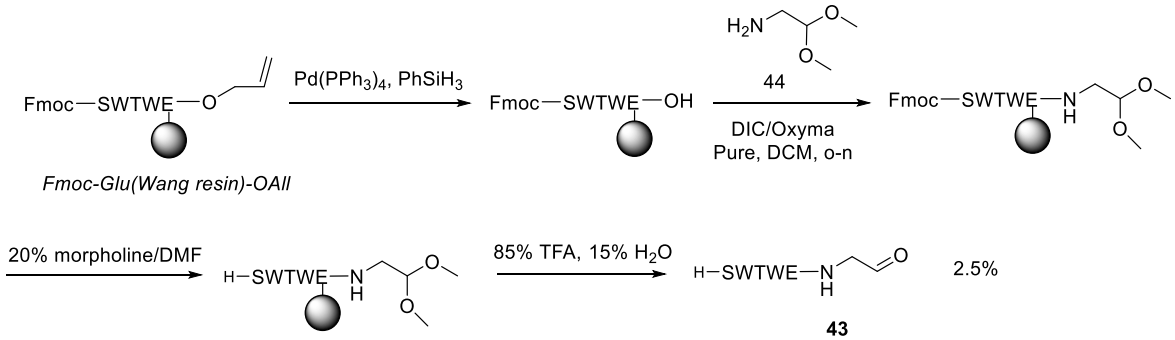
microwave-assisted SPPS to install the native amino acids followed by removal of the allyl protecting group.

Ethanolamine was initially considered as the starting material to install the aldehyde functionality (Scheme 63), but the material was volatile upon oxidation to aminoacetaldehyde and could not be isolated. In addition, aliphatic aldehydes are quite reactive, and therefore the acetal-protected **44** was used as the building block instead (Scheme 64).



Scheme 63. Initial strategy towards the coupling of the aldehyde moiety.

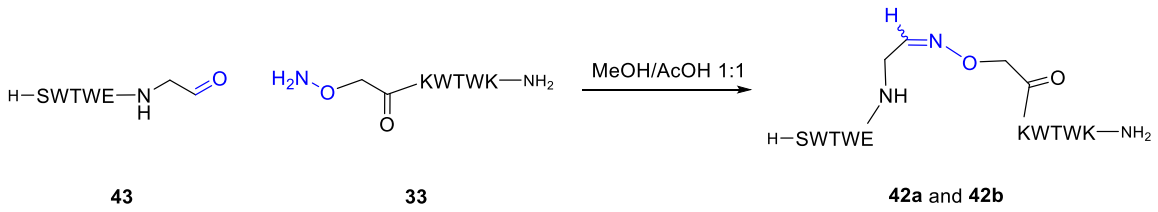
44 was initially coupled using DIC/Oxyma Pure in DMF, but deprotection of the *N*-terminal Fmoc group was observed. The coupling was therefore performed in DCM, with 10 equiv. of DIC and Oxyma Pure and 8 equiv. of **44** (Scheme 64). A test cleavage was performed, which showed complete conversion to the coupled product and maintenance of the Fmoc group through LCMS analysis. Interestingly, the product that was observed in the test cleavage was not the acetal-protected analogue but the deprotected aldehyde. Therefore, after Fmoc deprotection of the *N*-terminus, the cleavage of the peptide was performed in 85% TFA and 15% H_2O. This afforded the aldehyde-containing peptide **43** in a 2.5% yield after purification. The purification of this peptide was difficult, therefore ligation reactions were conducted with 83% pure material.



Scheme 64. Synthetic route towards 43.

5.7.2 Ligation towards 42a and 42b

The ligation between the two peptides **33** and **43** to afford **42a** and **42b** (Scheme 65) was performed using the conditions employed for the other mimics: at a peptide concentration of 1 mM, in AcOH/MeOH 1:1 and at room temperature. **42a** and **42b** have been named as such according to the order in which they are eluted from the column during RP-HPLC purification.



43 **33** **42a** and **42b**

Scheme 65. Ligation between 33 and 43 to generate 42a and 42b.

HPLC monitoring of the ligation step (Figure 158) showed complete product formation after 24 h. Although the starting materials did not reach complete consumption, no more product formation was observed after that time. After purification, **42a** and **42b** were isolated in a 54% total yield.

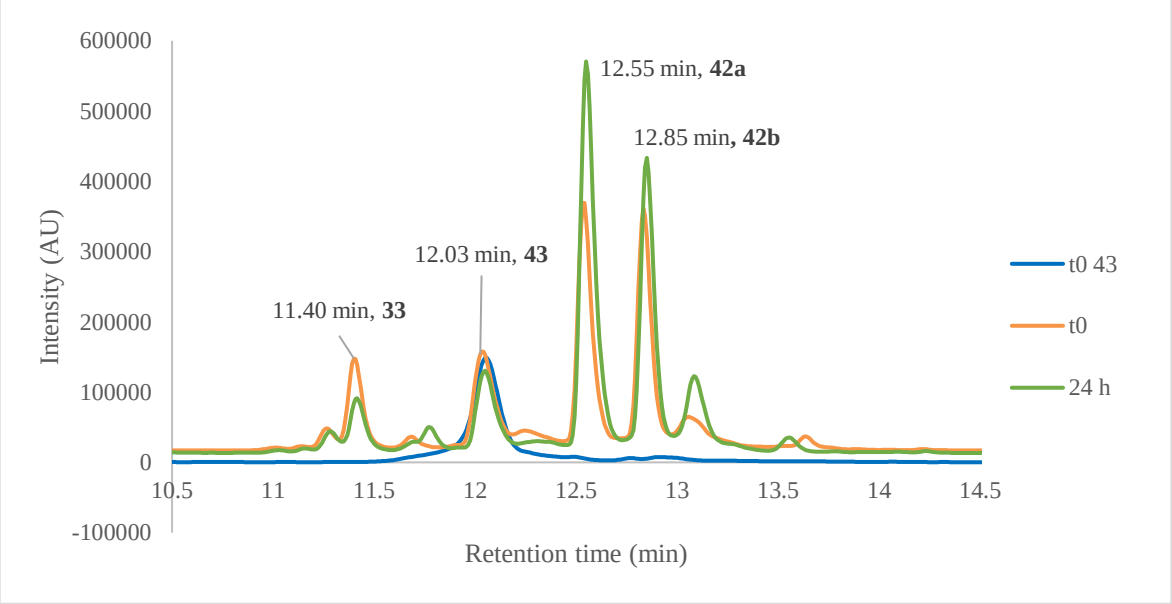


Figure 158. HPLC monitoring of the ligation towards 42a and 42b, showing three timepoints: t = 0 of 43 in solution before the addition of 33, t = 0 after the addition of 33 and after 24 h of reaction.

During RP-HPLC purification of the ligated products, interconversion between stereoisomers was observed. In an analogous manner to what was observed for **37a** and **37b**, acid-catalysed isomerisation of the oxime bond was likely taking place due to the 0.1% TFA

present in the solvent system. This isomerisation could take place through a hydrolysis-recondensation mechanism or the rotation of the C=N bond when in the iminium form (as discussed in Section 5.5.3).[344,345] In order to confirm that isomerisation was taking place, isomer interconversion tests were performed next.

5.7.3 Isomer interconversion tests

Isomer interconversion tests were performed for each stereoisomer in H_2O/MeCN 1:1 with 0.1% TFA, and the samples were monitored through HPLC. In the case of **42a** (Figure 159), interconversion was observed during the first day, after which time the proportion between stereoisomers was constant at a 70:30 ratio between **42a** and **42b**, respectively.

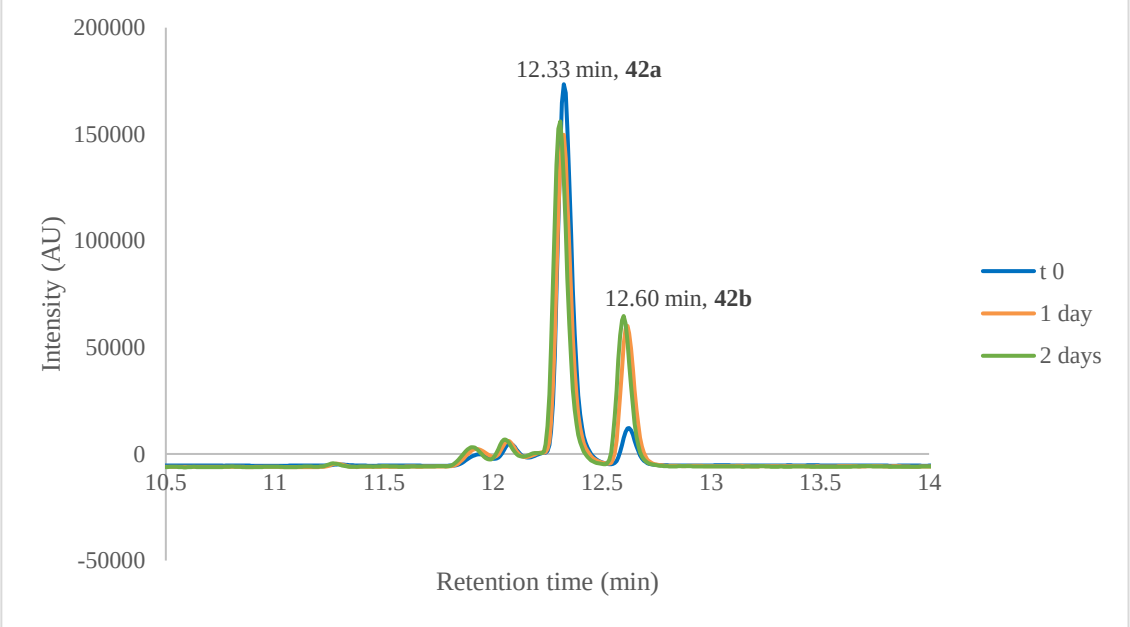


Figure 159. HPLC monitoring of the interconversion test for 42a.

Regarding **42b** (Figure 160), interconversion to **42a** was also observed during the first day and interestingly, the 70:30 proportion between stereoisomers was maintained after that time. This highlighted the higher thermodynamic stability of **42a**. In comparison to the other two oxime-linked mimics under study, **37a** and **37b**, this interconversion was extremely fast and afforded a ratio between stereoisomers instead of only one product.

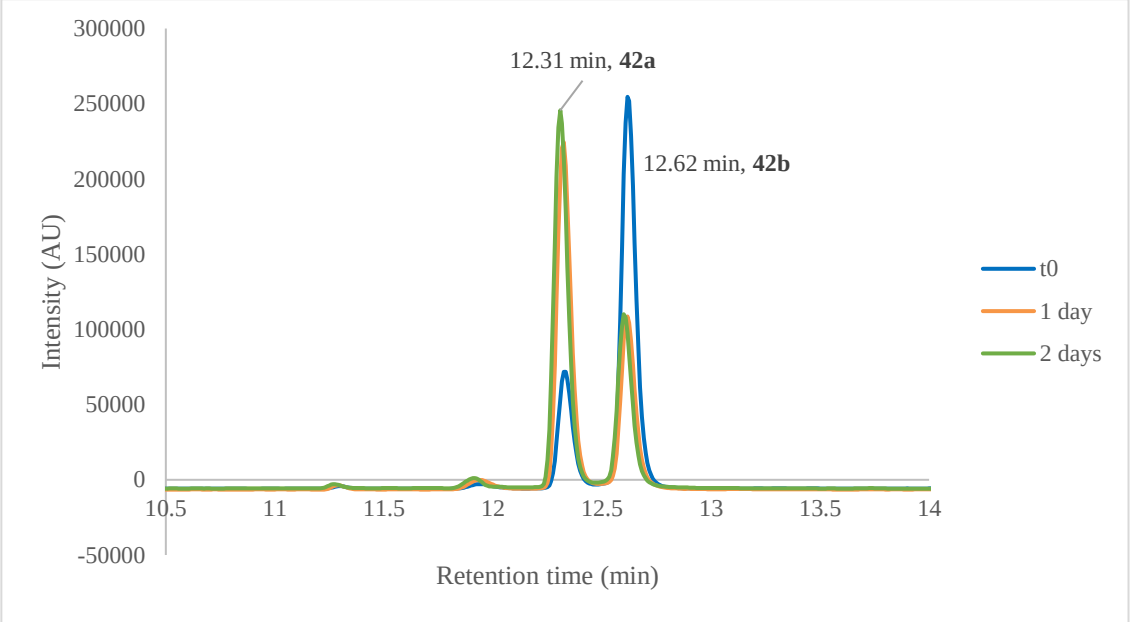


Figure 160. HPLC monitoring of the interconversion test for 42b.

5.7.4 Circular dichroism analysis of 42a and 42b

The high rate at which the stereoisomers interconverted caused the separation of both products to be difficult, hence the purity of the samples used for CD analysis was of 80% for each stereoisomer (with the remaining 20% corresponding to the other stereoisomer). CD experiments in the far UV were run for both analogues, and their folding was compared to that of the control sequence **TrpZip1**. A peptide concentration of 0.05 mg/mL (30 μM) was used, in a sodium phosphate buffer (20 mM) at pH 7 for **42a**. In the case of **42b**, when the same buffer was used precipitation of the sample took place. Therefore the experiments for this analogue were run in a HEPES buffer (10 mM) also at pH 7, under which conditions precipitation did not occur due to the lower ionic strength of this buffer. It was attempted to run experiments in the near UV using a 1 mg/mL concentration, but the low solubility of the conjugates did not permit this.

The far UV folding curves (Figure 161) showed that the two conjugates adopted β-hairpin conformations, with the characteristic minima at 212 nm and the maxima at 228 nm indicating β-strand structure and tryptophan stacking, respectively.[141,276] Although the difference in thermodynamic stability observed by HPLC analysis indicated a higher stability of **42a**, a significant difference in folding was not observed.

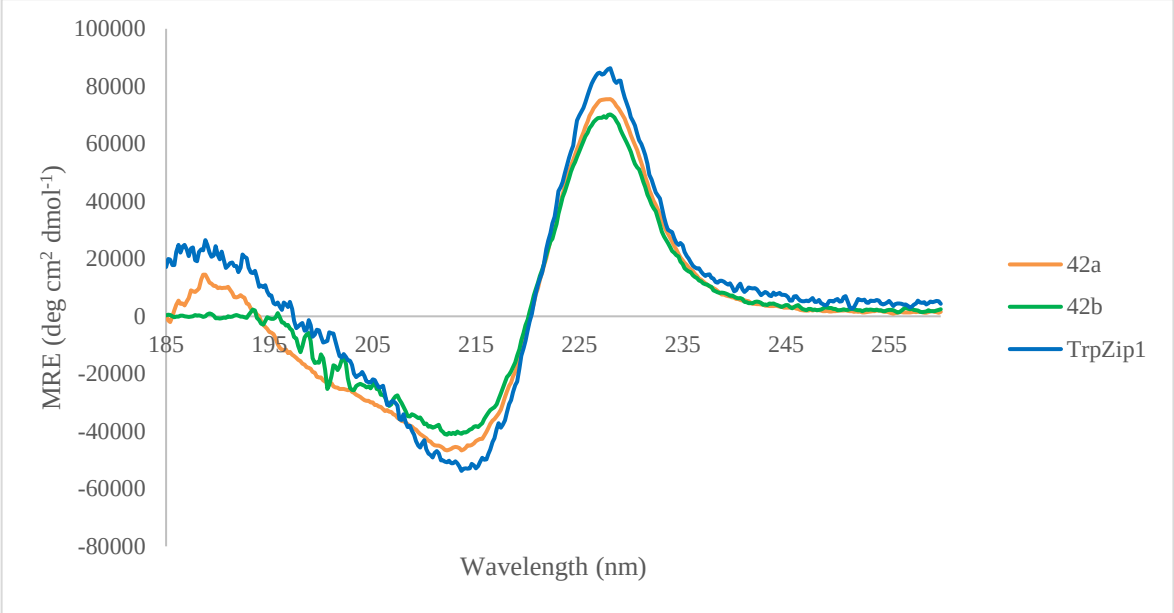


Figure 161. Far UV folding curves for the peptides in study.

In order to evaluate the thermal stability of the peptides, the change in MRE was monitored every 1 °C at 228 nm, from 5 to 80 °C. The folding curves before and after heating were superimposable for all analogues, indicating a reversible unfolding process. The denaturation curves obtained are presented in Figure 162.

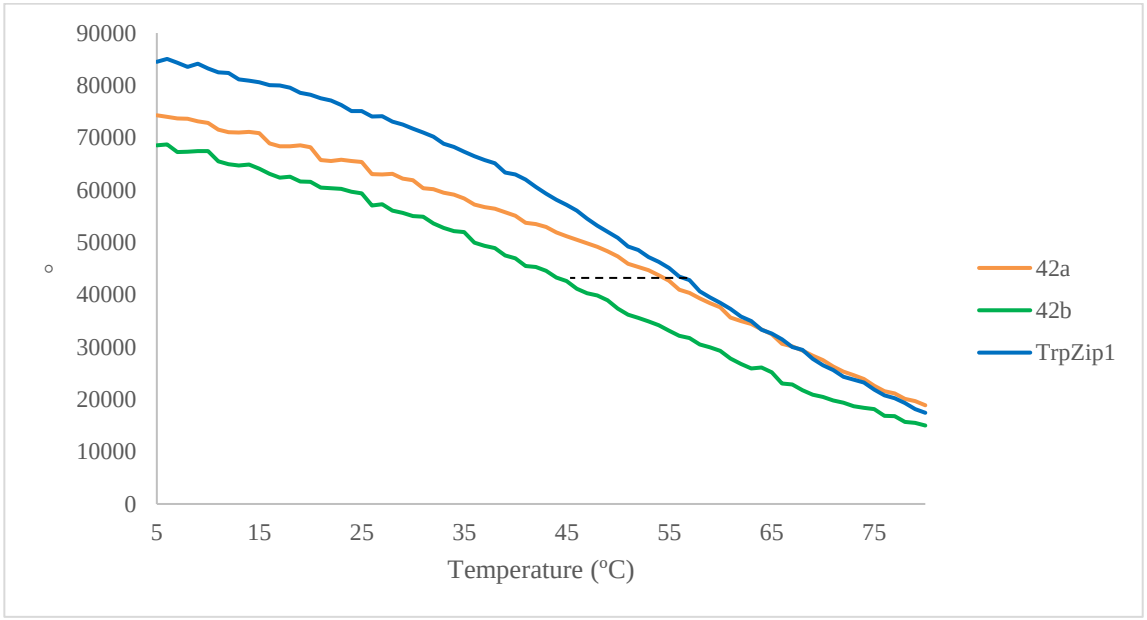


Figure 162. Thermal denaturation curves for the peptides in study. The black dashed line indicates an estimation of the midpoint of the curve for TrpZip1 and the point at which 42a and 42b adopt the same degree of folding.

Analogous to what was observed in the previous mimic system, **37a** and **37b**, the thermal denaturation curves of **42a** and **42b** showed a lower degree of folding in comparison to **TrpZip1** at equal temperatures. If the midpoint of the curve for **TrpZip1** is estimated, at 57

°C (black dashed line in Figure 162), the conjugates would adopt the same degree of folding at 55 °C for **42a** and at 45 °C for **42b**, 2 and 12 °C below the control, respectively.

As previously observed, a clear transition midpoint could not be obtained from the first or second derivatives of the denaturation curves (Figure 163). Therefore the fraction of folding calculation was used to obtain the melting temperatures.[277]

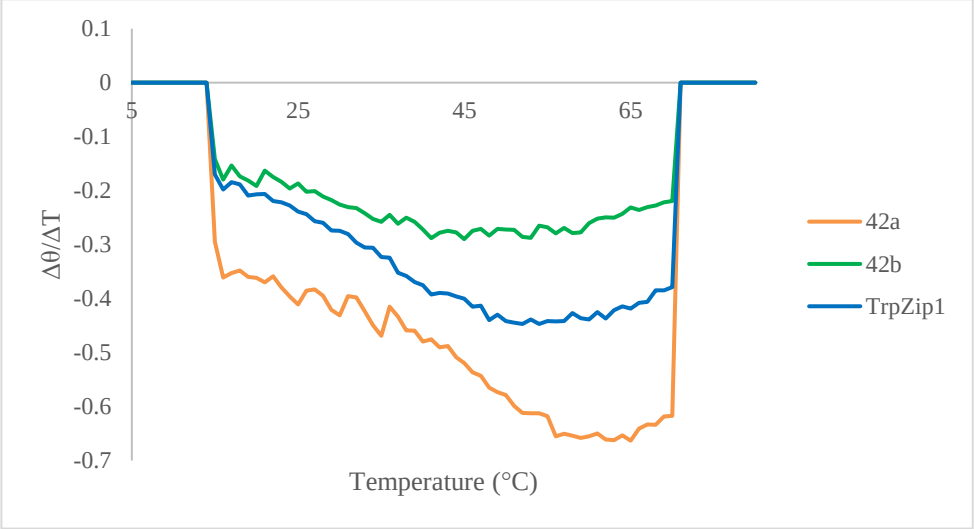


Figure 163. First derivatives of the melting curves for 42a and 42b.

As described for **37a** and **37b** (Section 5.5.4), the fraction of folding was calculated using an MRE of zero as the unfolding process endpoint and the MRE value of **TrpZip1** at 5 °C as the fully folded reference, due to the peptide conjugates being more unfolded than the control at that temperature. The curves obtained (Figure 164) showed a lower degree of folding of the conjugates in comparison to the control hairpin, with the three systems reaching the same unfolded point at 80 °C.

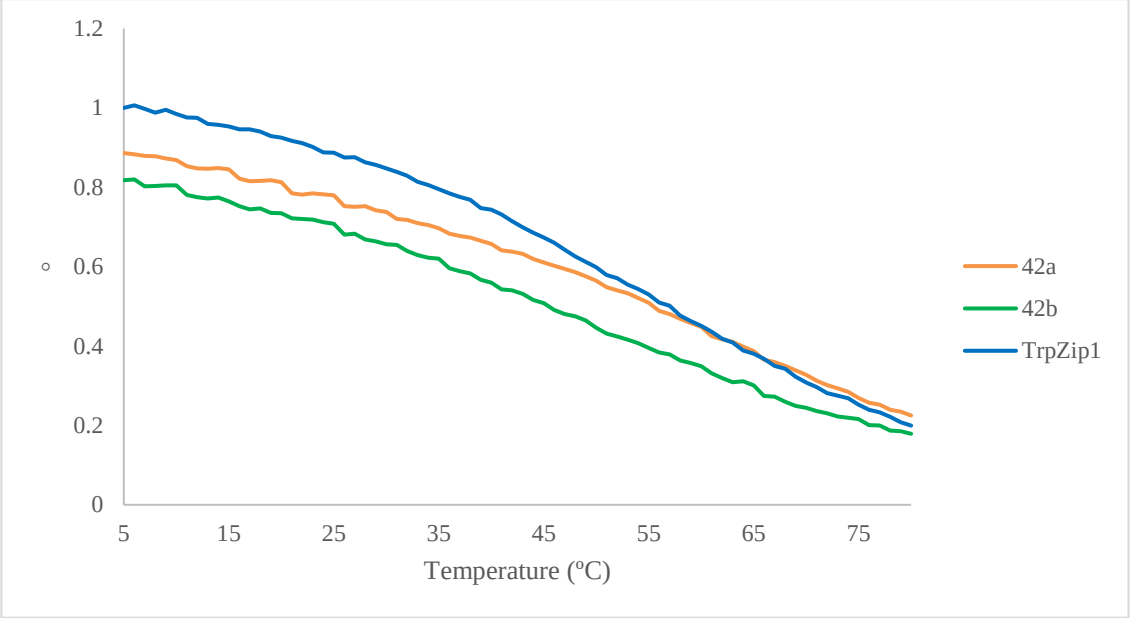


Figure 164. Fraction of folding curves relative to TrpZip1 for the peptides in study.

The T_m estimated for the systems in study (Table 24) showed that the melting temperature for **42a** was 2 °C lower than that of **TrpZip1**, whilst that of **42b** was lower by 12 °C. The difference in thermal stability between stereoisomers was relatively small, but it might indicate a correlation between the lower thermal stability of **42b** and the 70:30 ratio of **42a** to **42b** observed by HPLC analysis (Section 5.7.3). Comparing this information to the molecular dynamics simulations suggests that **42a** could correspond to the *Z*-oxime and **42b** to the *E*-oxime, as the latter was predicted to not establish the *i* to *i+3* like hydrogen bond and therefore would most likely be less folded than the *Z*-stereoisomer.

Table 24. T_m calculated for the peptides under study.

Sequence	T_m (°C)
TrpZip1	57
42a	55
42b	45

Due to the high rate of interconversion between the stereoisomers under some conditions, it was necessary to confirm whether isomerisation occurred under the conditions used for CD analysis (a neutral pH buffer and a high temperature). To this end, the samples used for CD analysis were analysed *via* HPLC before and after performing the thermal denaturation experiments (Figure 165). This showed that no interconversion had taken place.

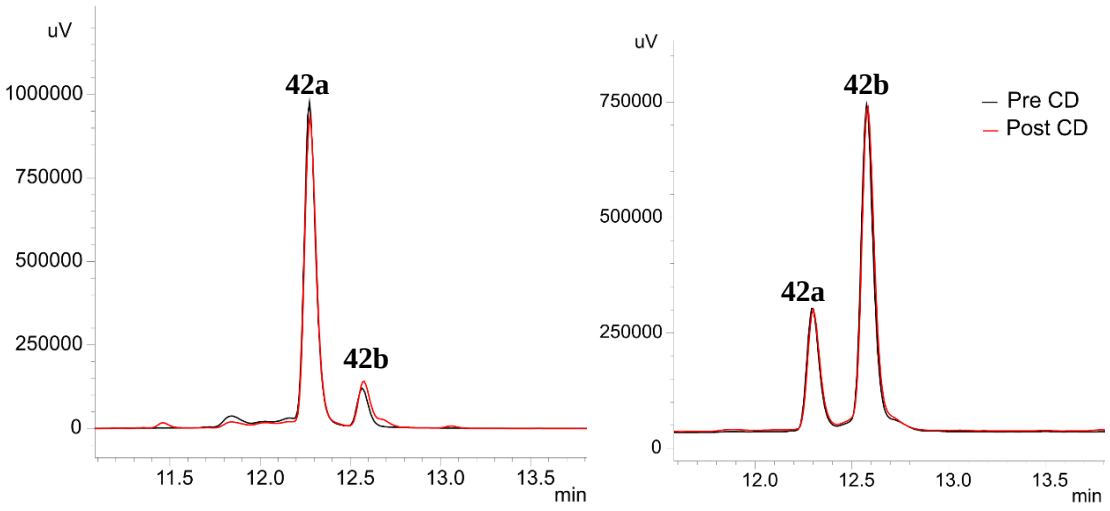


Figure 165. HPLC traces of the CD samples before and after the thermal denaturation experiments for 42a (left) and 42b (right).

In comparison to the *meta*-substituted aromatic ketoxime analogues (Figure 166), the folding of **42a** and **42b** was equivalent to that of **37a**, thus highlighting how a system with less preorganisation still favoured the adoption of a β-turn conformation. In addition the thermal stability was increased, from 35 °C for **37a** to 45 °C for **42b** and 55 °C for **42a**, possibly due to the higher constraints present in the aromatic ring-containing system.

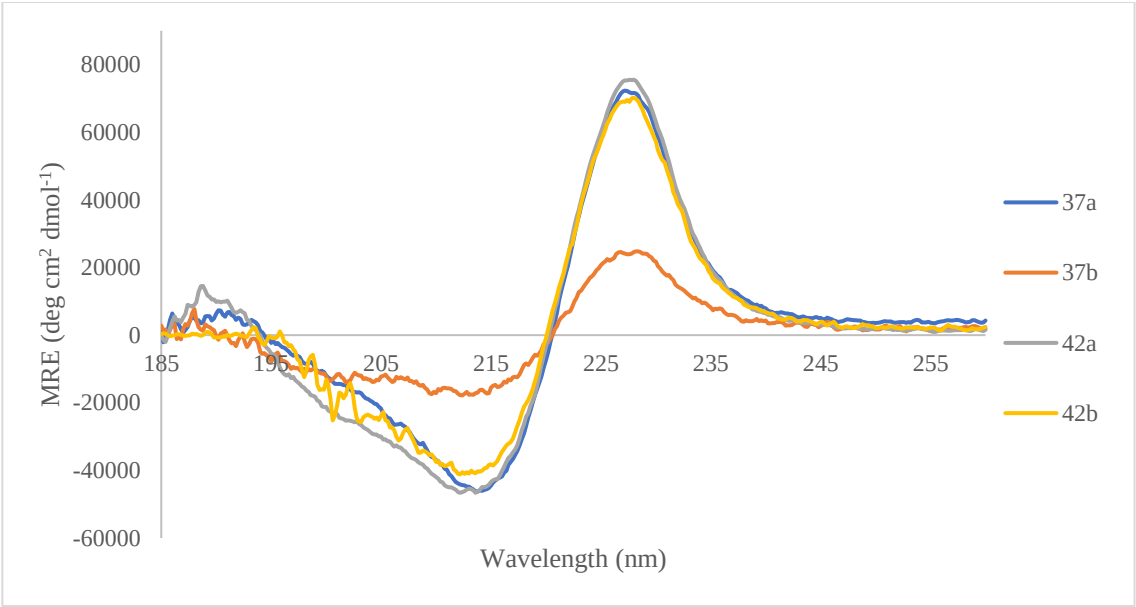


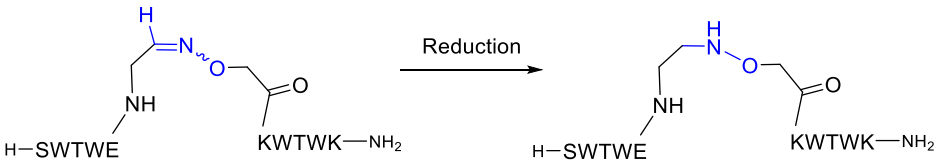
Figure 166. Comparison of the far UV folding between the *meta*-aromatic ketoxime analogues and the aliphatic aldoxime analogues.

All in all, the aliphatic conjugates designed adopted the same β-hairpin conformation as the control system and maintained an adequate stability. This showed how a β-turn mimic unit with less preorganisation could still favour β-hairpin formation and, in contrast to **37a** and

37b, the fewer constraints introduced by this system allowed both stereoisomers to adopt stable β-hairpins.

5.8 *N*-Alkoxy amine mimic

A mimic containing an *N*-alkoxy amine linkage was to be explored next, to ascertain whether the change to sp^3 centres at the ligation junction would have an effect in the folding and stability of the peptide. The aliphatic oxime-containing system, **42a** and **42b**, was chosen to perform this modification on. To this end, a reduction of the oxime bond would be performed (Scheme 66).



42a and/or **42b** **45**

Scheme 66. Reduction of 42a and/or 42b to furnish the *N*-alkoxy amine-bound mimic, 45.

Molecular dynamics simulations were run by Dr Drew Thomson, which showed that the *N*-alkoxy amine-bound conjugate maintained the β-hairpin conformation (Figure 167). The distance between the carbonyl oxygen and the amide proton in the *i* to *i+3* like hydrogen bond was extracted from the trajectory, which was found to be a mean of 1.9 Å. This suggested that the reduced linkage would establish a more tightly bound hairpin in comparison to its oxime-linked precursor (Figure 157).

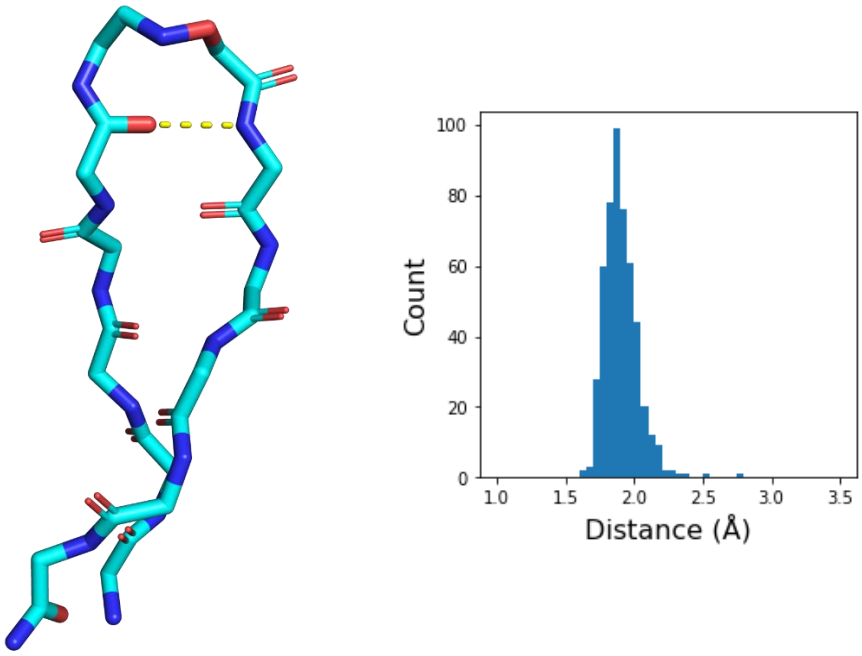
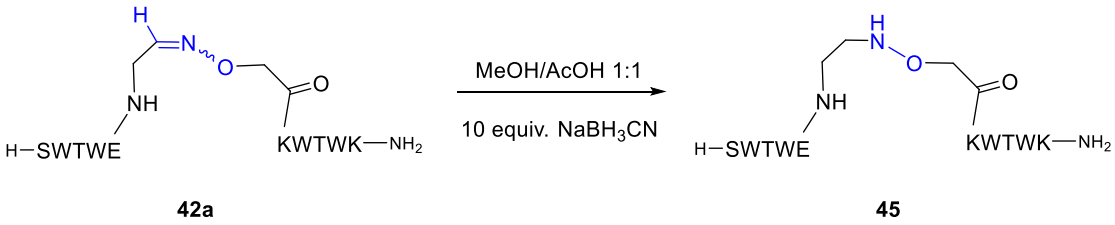


Figure 167. Left: representative frame extracted from the MD simulation of 45. Right: histogram of the distance between the oxygen in C=O(*i*) and the proton in NH(*i+3*).

5.8.1 Synthesis of 45

The reduction of the oxime bond to the *N*-alkoxy amine linkage was performed by dissolving a sample of **42a** (80% pure, with the remaining 20% corresponding to **42b**) in MeOH/AcOH 1:1, and adding 10 equiv. of sodium cyanoborohydride as the reducing agent (Scheme 67).



42a **45**

Scheme 67. Reduction of 42a to afford 45.

HPLC monitoring of the reduction reaction (Figure 168) showed that product formation was complete in 3 h, with 75% conversion of the starting material. Longer reaction times and addition of more reducing agent led to the formation of unidentified by-products instead of the desired compound. After purification, **45** was obtained in a 28% overall yield.

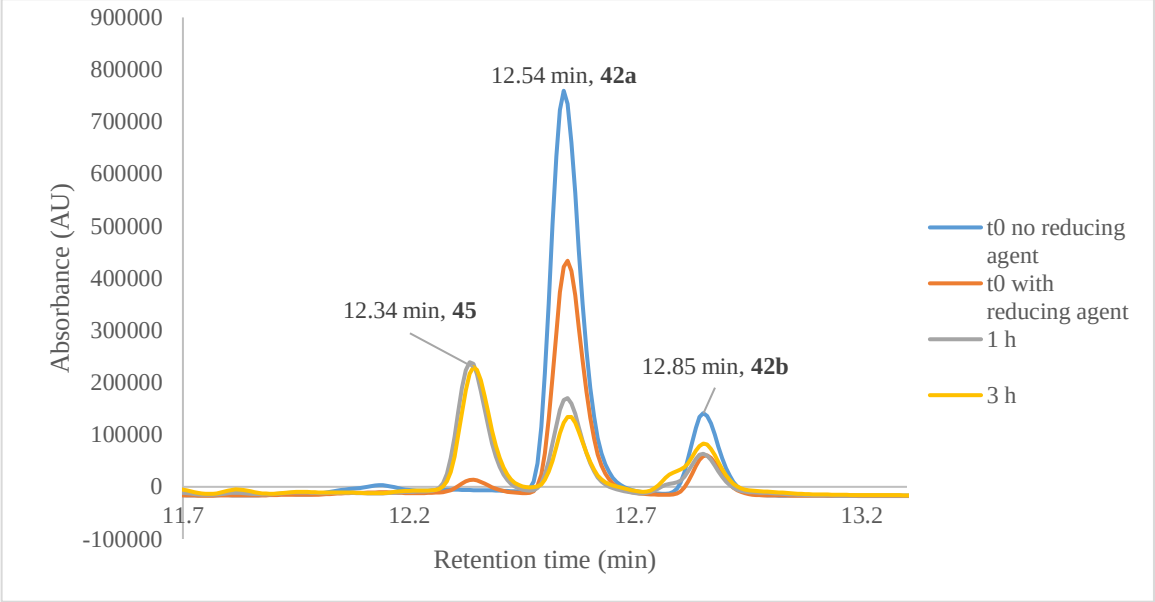


Figure 168. HPLC monitoring of the reduction towards 45.

5.8.2 Circular dichroism analysis of 45

The CD experiments for the *N*-alkoxy amine-linked conjugate were performed in the far UV at a 0.05 mg/mL concentration (24 μM), using a HEPES buffer (10 mM, pH 7) in order to prevent the solubility issues observed with the oxime-linked analogue **42b**. The far UV folding curve for **45** (Figure 169) showed successful mimicry of the β-turn, which gave rise to a well-folded β-hairpin showing the characteristic minimum and maximum. In comparison to **TrpZip1** and the analogous oxime-bound conjugates, the more pronounced peak and trough for **45** suggested it adopted a more tightly folded structure.

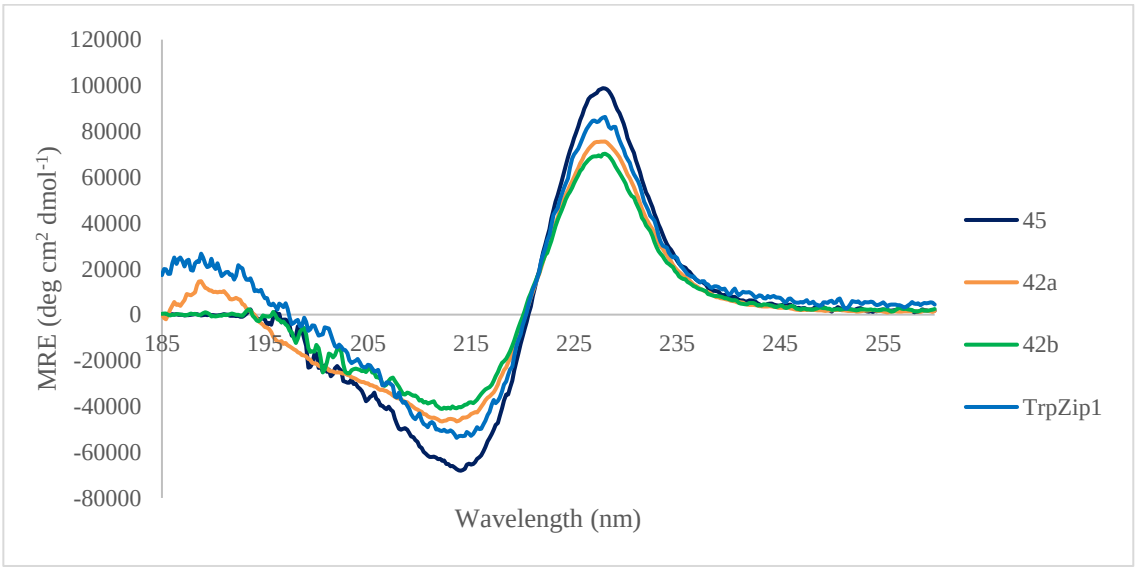


Figure 169. Far UV folding curves.

The folding of **45** was equivalent to that of **TrpZipGG** (Figure 170). The increase in folding observed when changing from an oxime to an *N*-alkoxy amine linkage in the conjugates, as well as from a Gly-Asn to a Gly-Gly turn in the controls, suggests that a higher flexibility in the β-turn results in a more stable hairpin structure for this system.

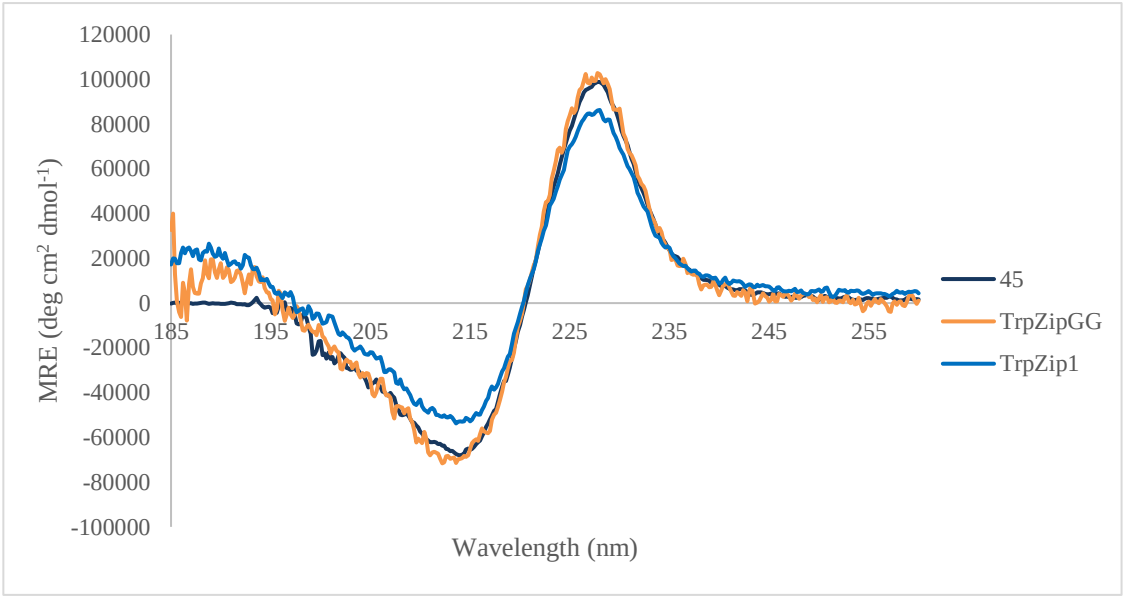


Figure 170. Far UV folding curves for 45 and both peptide controls.

The thermal denaturation curve at 228 nm was obtained for this analogue as described previously, and the folding curves in the far UV were superimposable after this process. The denaturation curve (Figure 171) had higher MRE values than those of the other conjugates, suggesting a higher stability, and its endpoint was at a higher point, indicating that a complete unfolding of the hairpin was not achieved in the temperature range used.

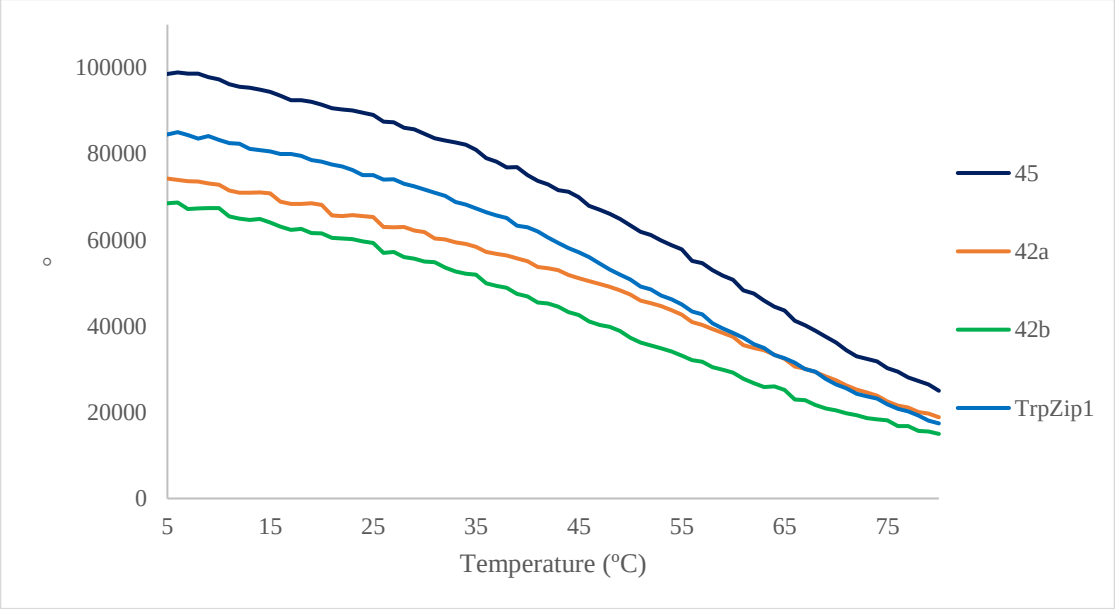


Figure 171. Far UV thermal denaturation curves for the peptides in study.

As observed with the other analogues under study, the derivative of the melting curve did not provide a clear transition midpoint (Figure 172).

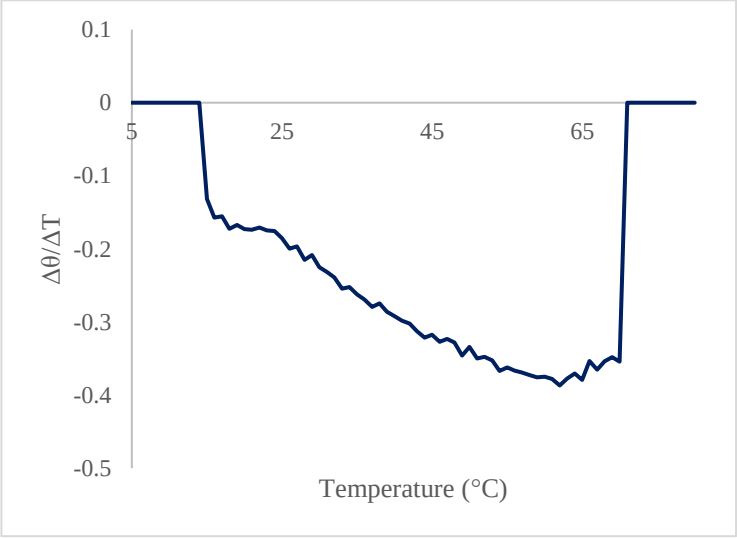


Figure 172. First derivative of the melting curve for 45.

In order to estimate the melting temperature for **45**, the fraction of folding was calculated for each temperature step. For this, the same strategy was used as described previously (Section 5.7.4), where **TrpZip1** was used as the folded reference and an MRE of zero was adopted as the unfolded value. These references were chosen for consistency, for the purpose of comparing the data. The curve obtained for **45** is presented in Figure 173, in comparison to those for **42a**, **42b** and **TrpZip1**. As it can be observed, the fraction folded for **45** at 5 °C

was higher than one, which indicated that the conjugate was more folded than the reference peptide.

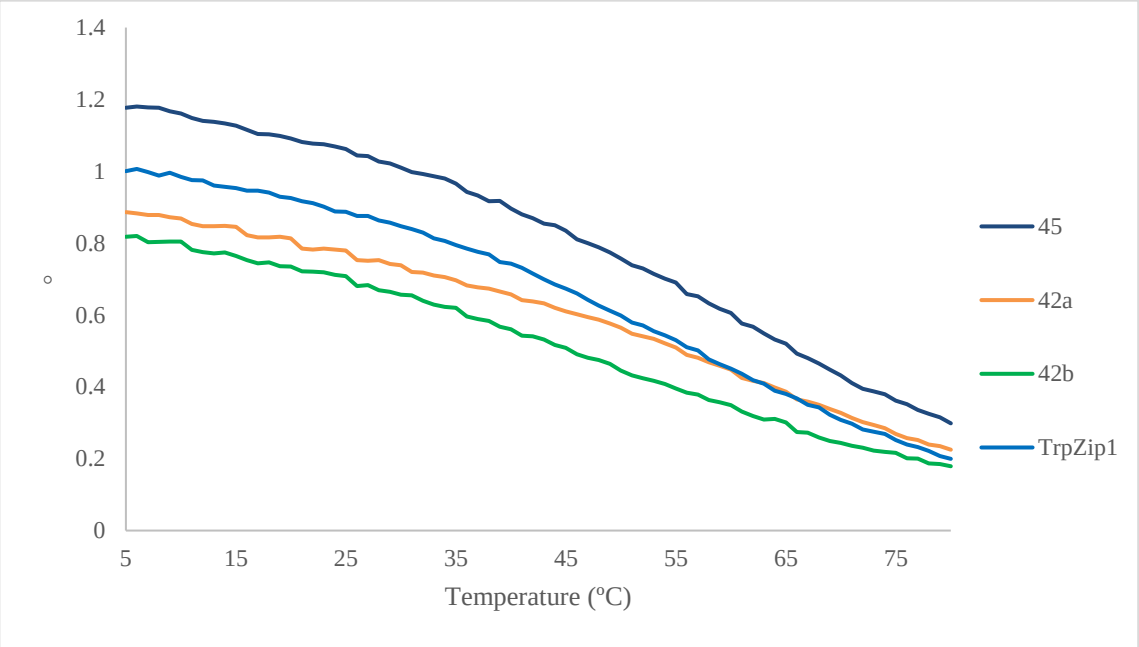


Figure 173. Fraction of folding curves relative to TrpZip1 obtained for the peptides in study.

The melting temperature estimated for **45** (Table 25) was the same as that of **TrpZipGG** and higher than that of **TrpZip1** by 9 °C. In comparison to its oxime-linked precursor, the T_m of **45** was higher than that of **42a** by 11 °C. The T_m obtained for these peptides are dependent on the references adopted, and as highlighted previously when **45** was analysed it was found that **TrpZip1**, adopted as the fully folded reference, was not fully folded. Changing this reference point would cause the T_m values to vary, but the relative behaviour of the peptides is still represented by the analysis with **TrpZip1** as the reference.

Table 25. Melting temperatures obtained from the fraction of folding curves.

Sequence	T_m (°C)
45	66
42a	55
42b	45
TrpZip1	57
TrpZipGG	66

The increased folding and thermal stability observed upon reduction to the *N*-alkoxy amine linkage could be due to the increased flexibility provided by the sp^3 centres in comparison to the sp^2 centres present in the oxime-bound conjugates.

5.9 Conclusions and future work

Novel oxime-linked β-turn mimic scaffolds have been designed. The first two targets used an aromatic ring to provide the right orientation for the reverse turn to take place. The *ortho*-substituted analogue **31**, analogous to the mimics explored in Chapter 3, could not be synthesised due to the low reactivity of the aromatic building block, which could not be coupled on the peptide fragment. Attention was thus shifted to a *meta*-substituted analogue, which showed successful mimicry of the β-turn. Two oxime stereoisomers were obtained which adopted different degrees of folding and thermal stabilities. **37a**, the *Z*-oxime, adopted a β-hairpin conformation analogous to that of the control whereas **37b**, the *E*-oxime, was largely unfolded.

Once an aromatic ring-based mimic was achieved, aliphatic oxime-based analogues were explored in order to evaluate if a less preorganised system would still favour the β-turn conformation. The ligation towards the initial target pursued, **39**, containing a phenyl substituent, was considerably slow. This was found to be due to the ketone-containing fragment and therefore the design was modified to contain an aldehyde moiety instead. The ligation towards **42a** and **42b** was significantly faster and afforded two peptides with different oxime stereoisomers that maintained the β-turn and β-hairpin conformation. In contrast to **37a** and **37b**, both aliphatic stereoisomers presented similar degrees of folding, and their thermal stabilities were 10 and 20 °C higher than that of the aromatic analogue **37a**.

An *N*-alkoxy amine-linked analogue **45** was explored, accessed by reduction of the oxime-bound **42a**. This peptide showed a higher degree of folding and stability in comparison to **TrpZip1** and its oxime precursors, likely due to the increased flexibility of the *N*-alkoxy amine centre. The folding of **45** was equivalent to that of **TrpZipGG**, showing that the folding and thermal stability is improved when reducing the oxime to an *N*-alkoxy amine linkage in the conjugates, and also when changing the native turn from Gly-Asn to Gly-Gly. This suggests that a higher flexibility in the β-turn provides a better folded hairpin conformation for this tryptophan zipper system.

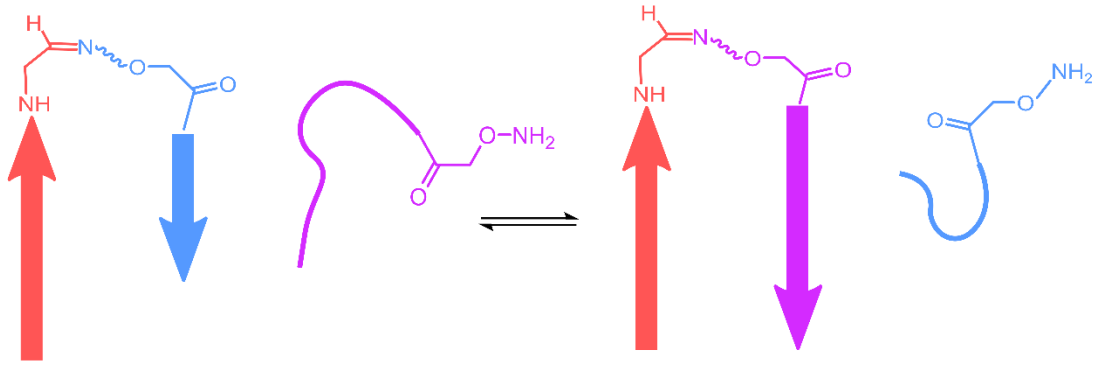
All in all, three novel β-turn mimic scaffolds have been designed and synthesised through a chemoselective oxime ligation. The mimics operate as designed and adopt a turn conformation that maintains the hydrogen-bonded β-hairpin structure observed in the control system. The comparison between an aromatic ring and an aliphatic oxime-containing scaffold indicated that systems with more conformational freedom can still favour a β-turn conformation, and the analysis of an *N*-alkoxy amine-bound analogue suggested that a higher linker flexibility could be correlated with better folding and thermal stability. This is a significant result, as it shows that opting for highly preorganised systems, although they provide the advantage of potentially inducing β-turn formation, is not always necessary. The remarkably high folding and stability of **45** opens the door to new possibilities in two ways. Firstly, a wide diversity of functionalities could be introduced within the β-turn scaffold in **45** that would enable the use of this system in a variety of contexts without altering the turn geometry (as described below). Secondly, other β-turn mimics analogous to **45** could be used within this system, and if the flexibility was maintained they would potentially still form the desired β-turn.

Future work could involve the introduction of additional functionality on the scaffolds presented in this chapter. For example, a different ketoxime structure could be pursued instead of the methyl-substituted **37a** and **37b** and the phenyl-substituted **39**, or additional substituents could be placed on the phenyl group of the latter in order to accelerate its ligation rate (this has been explored previously for aldoxime formation[194]). Another possibility would be the introduction of fluorescent groups on the oxime junction to be able to report on structure, function or binding to a target. In general, further functionalising the ketoxime linkages could be key in the development of new applications for these mimic systems.

Another strategy for future work could build on peptide **45** and focus on the study of the relationship between flexibility of the β-turn mimic and β-turn formation. Some possibilities that could be explored include modifying the number of methylene groups in the linker, altering the position of the oxime unit, or utilising a ketone instead of an aldehyde partner and investigating different substitutions.

The reversibility of the oxime linkage in acidic media also offers the opportunity to perform strand exchange experiments (Scheme 68). If a mixture of an oxime-bound peptide and a better suited *N*-alkoxy amine-containing peptide partner were allowed to reach thermodynamic equilibrium, the most stable product would ultimately be generated through

dynamic covalent exchange. This could find applications in dynamic covalent libraries,[314] where the pH could be tuned to initiate the exchange in an acidic medium, and stop it in a basic medium.[348] This control would allow for the isolation of the exchange products. Dynamic oxime exchange could also be exploited to identify novel PPI modulators by introducing the protein of interest as a template. The use of dynamic hydrazone exchange to generate novel peptide-small molecule hybrids that modulate PPIs has recently been reported.[349]



Scheme 68. Hypothetical strand exchange experiment. If an oxime-bound peptide with a suboptimal strand pairing and a better suited *N*-alkoxy amine-containing partner were mixed, the most stable product would ultimately be generated *via* exchange.

6 Conclusions

In this thesis, the control of peptide conformation has been explored through the design of conformational switching systems based on tertiary amide isomerisation as well as the mimicry of β-turns through chemical ligation.

Firstly, the study of **1** through a double mutant cycle (Figure 174) showed the presence of positive cooperativity between the Trp and Gln residues framing Pro. A sequence analogous to **1** with (*N*-Me)-Ala in place of Pro, **5**, was then studied through NMR. This showed a 14% increase in *cis* content, which could be caused by the establishment of a CH-π interaction between (*N*-Me)-Ala and Trp due to the high flexibility of the former.

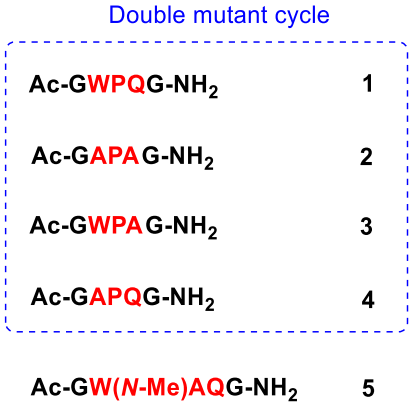


Figure 174. Double mutant cycle peptides (in blue box) and *N*-methylated Ala sequence studied in Chapter 2.

Disulfide-actuated switches were then designed using this information (Figure 175). The first generation was based on a CxPC sequence, which explored Trp (**6**) and Ala (**7**) preceding Pro. Conditions for the disulfide formation were sought, which identified the use of DPDS as the most suitable strategy. The study of the disulfide-bridged peptides **8** and **9** through NMR did not show a switching behaviour, with an increase of 3-4% *cis* Pro being observed upon oxidation. The second generation of switches made use of vicinal disulfide bridges and explored both Pro (**10**) and (*N*-Me)-Ala (**11**). Although a 10% increase in *cis* configuration was observed when substituting Pro by (*N*-Me)-Ala, the systems did not behave as a switch upon oxidation and instead adopted up to five different conformations. This could be due to the *C*-terminal Cys adopting a *cis* configuration and the presence of two disulfide bridge rotamers.

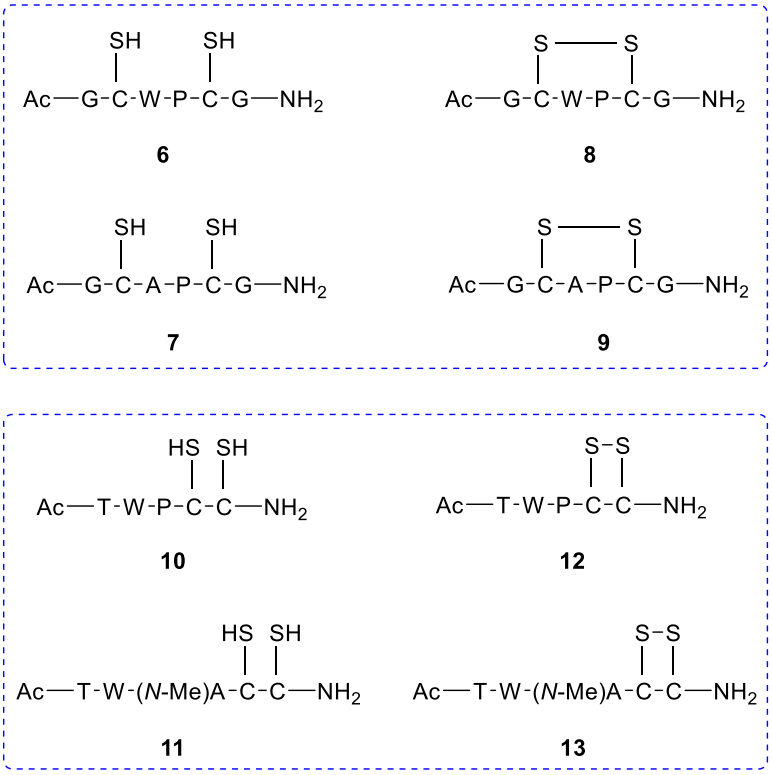


Figure 175. Disulfide-actuated switch designs explored in Chapter 2.

In Chapter 3, novel β-turn mimics installed through a reductive amination step were studied. The Tryptophan Zipper **TrpZip1** was used as the control system, where the mimics would be installed instead of the Gly-Asn turn. Two β-turn mimic scaffolds were designed, based on an *ortho*-substituted aromatic ring, and they were implemented within tripeptide backbones as well as the full-sized TrpZip β-hairpin. A reductive amination protocol was developed to ligate two unprotected peptide fragments, which utilised sodium cyanoborohydride in MeOH/AcOH 1:1 and provided full conversion to the ligated peptides in 15 min. The tripeptide systems **19**, **20a** and **20b** were analysed through NMR, which showed that a turn-like conformation was adopted in solution. The β-hairpin systems **14**, **(*S*)-15** and **(*R*)-15** (Figure 176) were evaluated through CD and NMR. CD analysis showed that **14** and **(*S*)-15** had an increased thermal stability in comparison to **TrpZip1**, and that **(*S*)-15** and **(*R*)-15** had different stabilities, with **(*S*)-15** being the more stably folded diastereomer. NMR analysis showed successful mimicry of the β-turn, with the *i* to *i+3* like hydrogen bond being present in the three conjugates, and maintenance of the hydrogen bonded hairpin as well as of the side chain conformations. The diastereomers in **(*S*)-15** and **(*R*)-15** were identified by comparing the strength of the NOE cross peaks to the distances extracted from MD simulations.

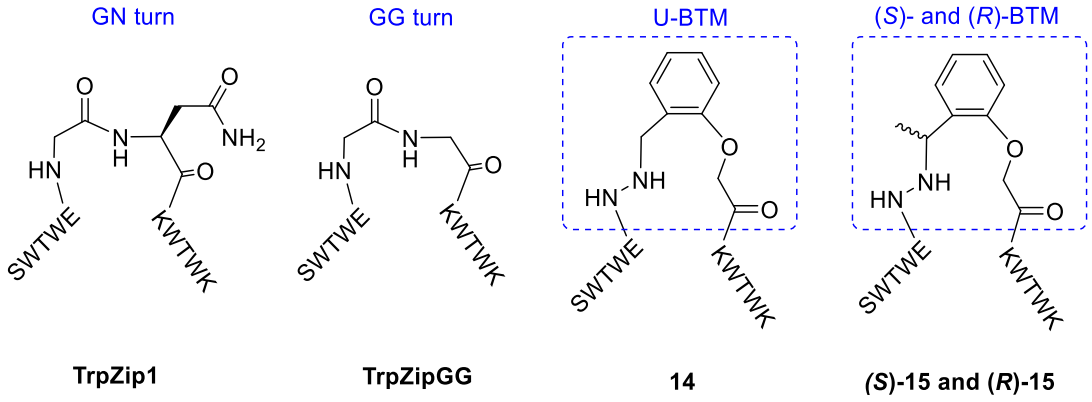


Figure 176. Original TrpZip1 sequence, TrpZipGG and unnatural β-turn units explored in Chapter 3.

The aim of Chapter 4 was to design a cyclic inhibitor of the MDM2/p53 interaction. To this end, the (D-Pro)-(L-Pro) turn in **29** was substituted by the U-BTM β-turn mimic designed in Chapter 3, giving rise to the conjugate **28** (Figure 177). The control sequence was accessed using an on-resin head to tail cyclisation procedure. The conjugate **29** was synthesised using analogous conditions to the ones developed in Chapter 3, which afforded the hydrazone-linked cyclic peptide **30** upon cleavage. Reduction to access the hydrazide-bound **28** was performed using sodium cyanoborohydride. Initial SPR assays were performed, which showed a two-fold improvement in the binding affinity of the peptides upon introduction of the non-natural β-turn. This was postulated to be a result of the higher flexibility of the mimic peptide in comparison to the (D-Pro)-(L-Pro) turn, or due to the establishment of additional interactions between the MDM2 binding pocket and the β-turn scaffold. These results were positive yet preliminary, and would need to be confirmed by repeating the SPR assays with a more stable protein construct.

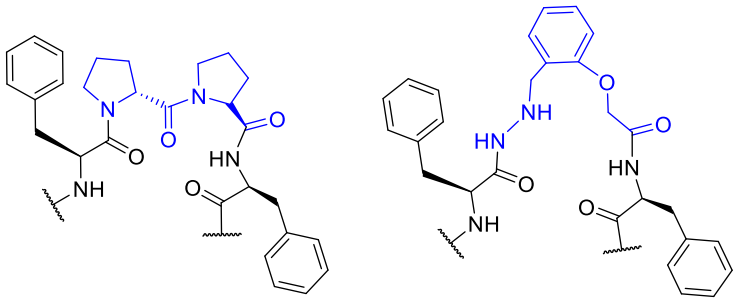


Figure 177. Control sequence (left) and conjugate (right) studied in Chapter 4.

Lastly, oxime-bound β-turn mimics were designed (Figure 178), which were implemented within the TrpZip β-hairpin. The initial target, **31**, was a mimic based on an *ortho*-substituted aromatic ring. This system could not be accessed due to the low reactivity of the aromatic building block, which prevented the synthesis of the *N*-terminal fragment **32**. The design was then adapted to a *meta*-substituted aromatic ring, **37a** and **37b**. Ligation conditions were investigated, and it was found that employing MeOH/AcOH 1:1 provided the fastest reaction rate and a higher yield. CD analysis of both stereoisomers showed that **37a** successfully mimicked the turn and hairpin structure present in the control, while **37b** was largely unfolded. By comparison to MD simulations, **37a** was proposed to be the *Z*-oxime and **37b** the *E*-oxime. A further design was then explored, containing a phenyl-substituted oxime linkage. Although the peptide fragments could be synthesised, the ligation to access **39** was extremely slow due to the ketone-containing peptide fragment **40**, and therefore the design was modified. An aliphatic oxime linkage was then investigated (**42a** and **42b**), and both stereoisomers were found to mimic the β-hairpin conformation successfully by CD. By comparison to the MD models, **42a** was proposed to be the *Z*-oxime and **42b** the *E*-oxime. The aliphatic oxime system was reduced, and the *N*-alkoxy amine-bound analogue **45** was studied through CD. This showed that **45** had a higher degree of folding and stability in comparison to the other conjugates studied in this chapter, as well as in comparison to the native control. These studies suggested that in this system, less preorganised mimics can still replicate the β-turn conformation, and that an increased linker flexibility provides a higher degree of folding and stability.

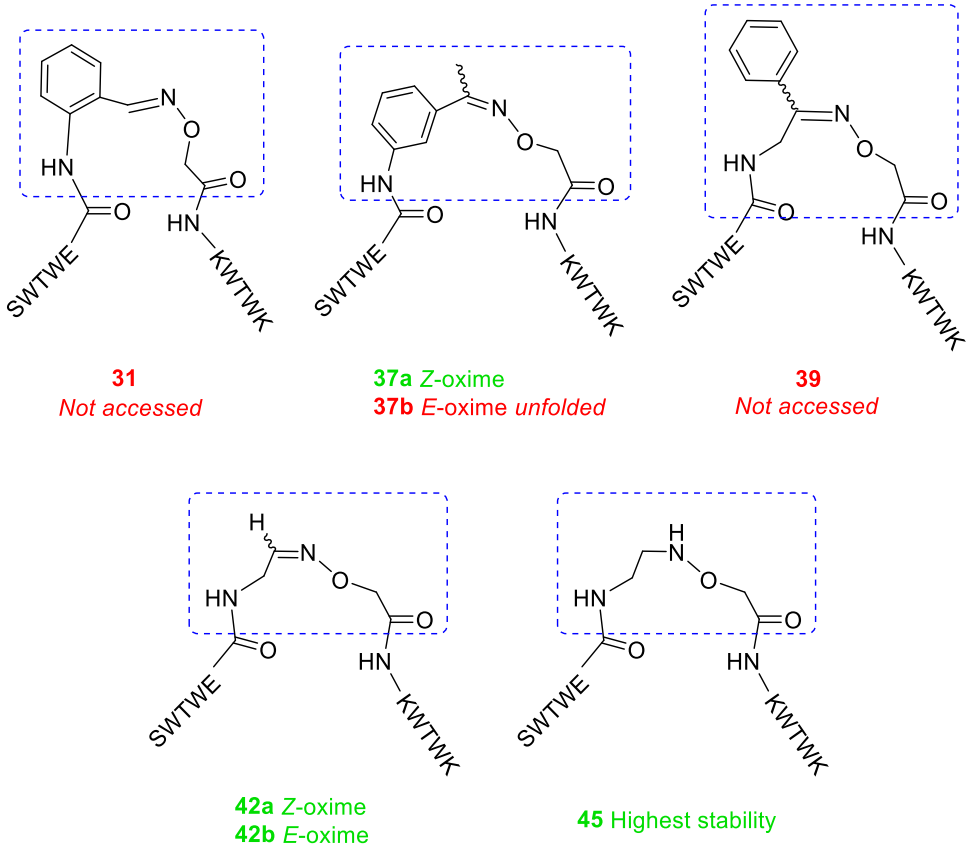


Figure 178. Oxime-bound β-turn mimics studied in Chapter 5.

In comparison to the hydrazide-linked mimics introduced in Chapter 3 (Figure 179), the *meta*-substituted **37a** and the aliphatic oxime system **42a** and **42b** shared equivalent degrees of folding, which were slightly lower than that of **14** and the control **TrpZip1**. Their thermal stabilities in comparison to **TrpZip1** were also lower, with the preorganised **37a** being the least stable conjugate. This highlights how in this case, the *meta*-substituted aromatic ring did not favour the turn conformation as much as the *ortho*- system used in **14**. In contrast, the *N*-alkoxy amine-bound **45** was the most folded analogue and it had one of the highest thermal stabilities of all peptides under study, 66 °C. This was only surpassed by the hydrazide-linked conjugate **(*S*)-15** with a T_m of 70 °C. This showed how the flexibility of the aliphatic *N*-alkoxy amine linker provided a similar folding and stability as the (*S*)-diastereomer in the more constrained conjugate **(*S*)-15** for this hairpin system. In comparison to **14**, it favoured a more suitable turn geometry than the preorganised *ortho*-substituted aromatic ring.

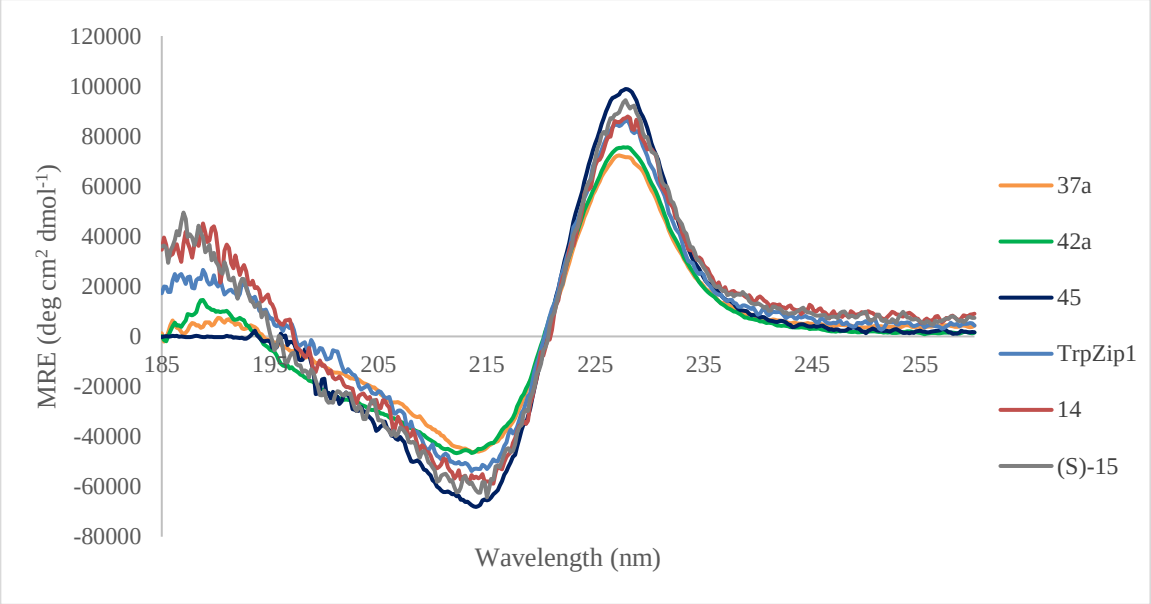


Figure 179. Comparison of the folding in the far UV for the mimics under study.

Overall, in this work the control of peptide conformation has been explored *via* different approaches. Tertiary amide isomerisation has been shown to be an attractive backbone onto which switching systems can be designed, with (*N*-Me)-Ala providing higher *cis* contents in comparison to Pro. The cooperativity between residues framing the tertiary amide has been confirmed, and the design of disulfide-based switches has been attempted. Thus, this work has set the foundations for new and improved switches to be designed. Building from the non-vicinal disulfide systems, (*N*-Me)-Ala could be exploited in order to facilitate the access to higher *cis* tertiary amide contents, and different disulfide placements could also be investigated. Inspiration could be drawn from β-hairpins and β-turns, as disulfide bridges have been found to stabilise hairpins when placed at non-hydrogen bonded facing positions.[126] The vicinal disulfide systems showed greater complexity than initially expected, but this behaviour clearly indicates the substantial influence of this constraint on peptide conformation. This suggests that this disulfide connectivity, although intricate, has great potential that could be exploited for the purpose of designing a conformational switch.

Novel β-turn mimics have been designed, accessed through a reductive amination and an oxime ligation step, which closely mimic the conformation of the control β-hairpin. This work provides the ability to install a β-turn mimic at the same time as ligating two unprotected peptide fragments, using either a reversible or non-reversible linkage, which is a tool that could have wide applications in chemical biology. Furthermore, the wide range of β-turn units explored compose a toolkit that could find applications in diverse systems, as they provide different degrees of folding and stability, functionalisation, flexibility, and

installation methods. These β-turn mimics could be used in the synthesis of miniproteins, and their further functionalisation could enable the discovery of novel PPI inhibitors. Regarding the latter, groups that would improve the affinity to a desired target could be added, as well as a poly-Arg chain to improve the cell-penetrating properties of the peptide. Fluorescent groups like a FRET donor or acceptor could be inserted in these units in order to report on binding and function. Other fluorescent groups could be introduced to facilitate live cell imaging, as many of the fluorophores used for this purpose require a coupling step to one of the termini of the peptide or a side chain. This would be particularly useful in the case of cyclic peptides, in order to preserve the side chain functionality. Finally, and as previously discussed, the reversibility of the oxime linkage could be further explored to facilitate its use in dynamic combinatorial libraries.

7 Experimental

7.1 Database

The database was constructed by Dr Drew Thomson using the PISCES server to identify a subset of high-resolution, non-sequence redundant protein crystal structures (resolution 2.5 Å or better, R-factor 1.0 or better, PDB version 28/10/2019).

A region of secondary structure was defined as having four or more contiguous residues of the same secondary structural assignment by DSSP.[350] Following this definition, a 'loop' was defined as unstructured by DSSP, but they were allowed to contain up to three contiguous residues with the same DSSP secondary structural assignment, as well as blocks of mixed secondary structures. The condition that a loop needed to be flanked on each side by a region of secondary structure was applied, therefore excluding *N*- and *C*-terminal unstructured regions. Loops were further checked to exclude any in which the loop was discontinuous due to missing residues or atoms in the crystal structure. For each of the loops thus identified, a database entry was generated containing the 3D coordinates for the loop, as well as the four residues of secondary structure on either side. The database entry also contained information such as the sequence, DSSP assignment, and vectors representing the end-to-end separation and orientation of flanking secondary structure for each loop.

7.1.1 Search for Xaa-Pro-Yaa motifs

In order to identify the influence of neighbouring sequence on the formation of *cis* proline residues, the database was searched for all 400 possible sequences based on the Xaa-Pro-Yaa pattern. For each subsequence, the proline was assigned as *cis* if the ω dihedral angle for the amide bond was 0 +/- 20°, and *trans* if the angle was 180 +/- 20°. Entries were discarded for a small number of sequences for which ω fell outside either of these ranges. For each subsequence, the percentage of *cis* was calculated.

7.2 Energy minimisation models

Energy minimisation models were built on Avogadro[351] software, using the MMFF94 forcefield and a conjugate gradients method until the energy value was stabilised.

7.3 Molecular dynamics simulations

7.3.1 Chapter 2

Simulations were run by Miss Bethany Atkinson using Chimera, and the molecular dynamics simulations were carried out using Gromacs 4[352] with the Plumed 2.5 plugin.[353] *N*-terminal acetyl and *C*-terminal amide capping groups were used. The peptides were immersed in a cubic box of explicit water with a minimum of 1.0 nm between the peptide and the edge of the box. A steepest descent algorithm was then used, followed by a four-stage equilibration process. First a 50 ps NVT ensemble was carried out followed by a 50 ps NPT ensemble both with restraints of 1000 kJ/mol. A second NVT ensemble was then carried out for 100ps without restraints. Finally, a 100 ps NPT ensemble without restraints was carried out and the trajectory file was then used in the metadynamics simulation. A separate V-rescale thermostat was used for both the peptide and the solvent. A Berendsen barostat with an isothermal compressibility of 4.5 x 10^{-5} bar^{-1} was used for the NPT ensembles and the metadynamics simulations. An AMBER-99SB-ILDN force field with tip3p water was used and the simulations were carried out at 300 K and 1 bar. For the duration of the simulation all bonds to hydrogen atoms were constrained to equilibrium values using a LINCS algorithm. Nonbonded interactions (both Lennard-Jones and electrostatic) were truncated at 1 nm. For long range electrostatics a particle mesh Ewald (PME) with an order of 4 and Fourier spacing of 0.12 nm was used. The metadynamics simulations were carried out with two collective variables: the ψ dihedral angle of the proline in the peptide, and the improper dihedral angle ζ. The simulations were carried out with a timestep of 1 fs using a leapfrog algorithm for 100 ns. Gaussian hills of height 0.2 kJ/mol and a sigma value of 0.2 radians for ψ and 0.15 radians for ζ, were added every 1 ps.

To estimate the energy difference between the *cis* and *trans* peptide in solution the ψ value was integrated out and the free energy as a function of ζ was plotted every 20 ps. An average over multiple of these fes files once convergence was reached gave the most accurate estimate of the free energy surface of the peptide with changing ζ.[354] The difference between these two minima was then used as the ΔG value between *cis* and *trans* proline in the peptide. To obtain the ΔG value between the *cis* and *trans* proline, the minimum point between -20 to 20° was taken as the *cis* value and the minimum point between below -160 or above 160° was taken as the *trans* minimum. The difference between the two was then calculated by taking the *trans* minimum minus the *cis* minimum. The difference between the minima

energies was then plotted against the time in the simulation which the fes file was taken from. After the simulation has reached convergence, the average over all the points plotted can be found and used as the ΔG value between the *cis* and *trans* Pro states. To determine the point at which to start averaging the values, the minimum energy level (E_{min}) throughout the simulation and the value closest to half this minimum value (1/2 E_{min}) were found (Figure 180). The time difference between these points was multiplied by 3. Three lengths of time after ½ E_{min} was used as the starting time point for averaging the ΔG values.

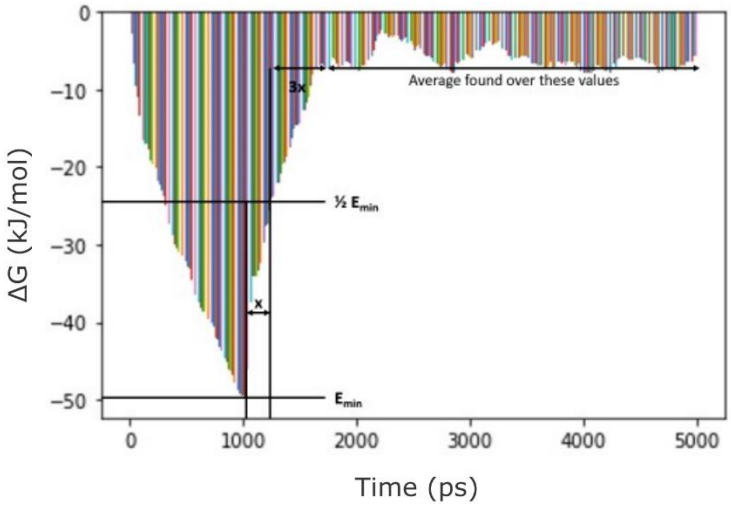


Figure 180. Determination of the point at which the MD simulation for 2 reached convergence.

7.3.1.1 Simulation of 1

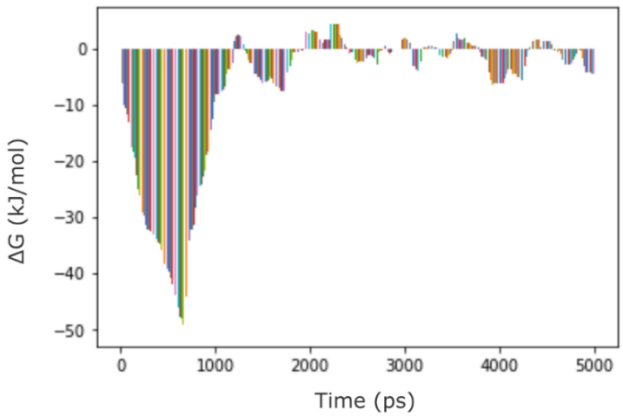


Figure 181. ΔG between the *cis* and *trans* states along the simulation.

7.3.1.2 Simulation of 2

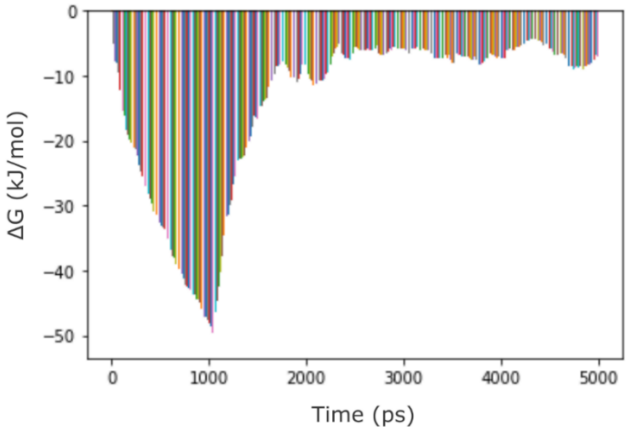


Figure 182. ΔG between the *cis* and *trans* states along the simulation.

7.3.1.3 Simulation of 3

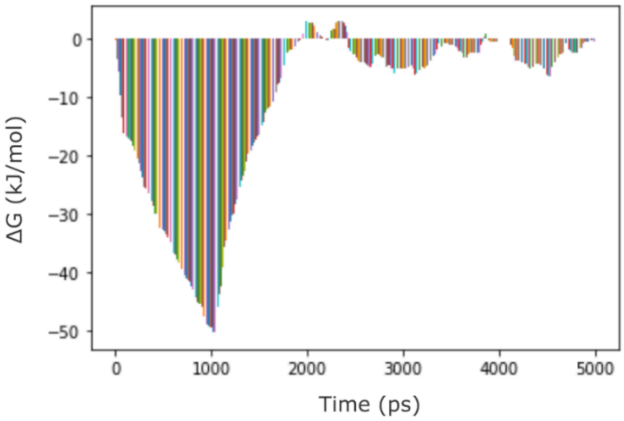


Figure 183. ΔG between the *cis* and *trans* states along the simulation.

7.3.1.4 Simulation of 4

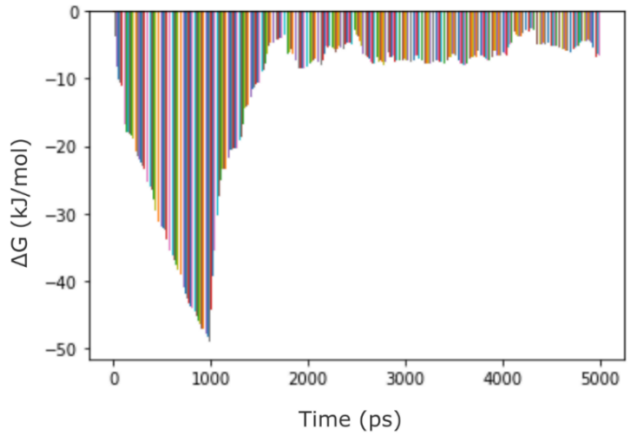
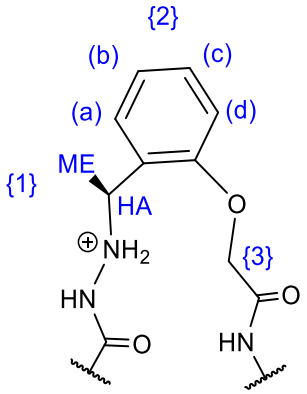


Figure 184. ΔG between the *cis* and *trans* states along the simulation.

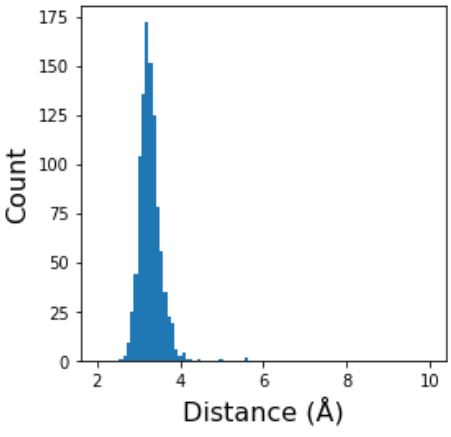
7.3.2 Chapters 3 and 5

Molecular dynamics simulations were run by Dr Drew Thomson. The simulations were set up using AmberTools[355] and run using OpenMM.[356] Fragment structures for the β-turn mimic subunits were parameterised using AmberTools, and starting models were built using Avogadro.[351] Simulations were set up with a cubic water box, and with compensating sodium and chloride ions such that the simulation was charge neutral. Periodic boundary conditions were used, and a Monte-Carlo barostat was employed. Simulations were run at 1 atm pressure and at a temperature of 300 K. Simulations were run with 2 ps step size for a total of 100 ns. The amber ff14SB forcefield was used. Trajectories were analysed using the MDanalysis python module.[357]

7.3.2.1 Distances extracted from the simulation of (*S*)-15

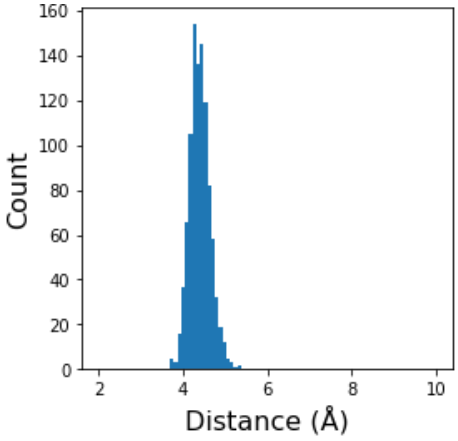


{1}ME (carbon) – {3}*pro*-R



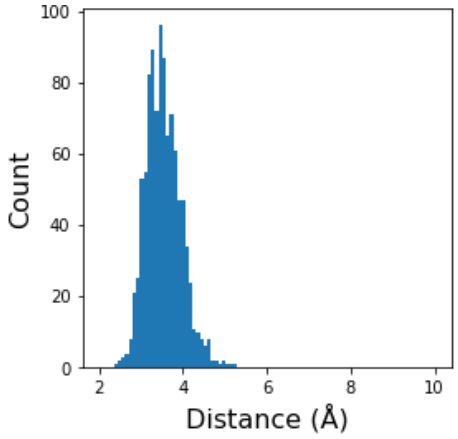
Average 3.3, min 2.6, max 5.6 Å

{1}ME (carbon) – {3}*pro*-S



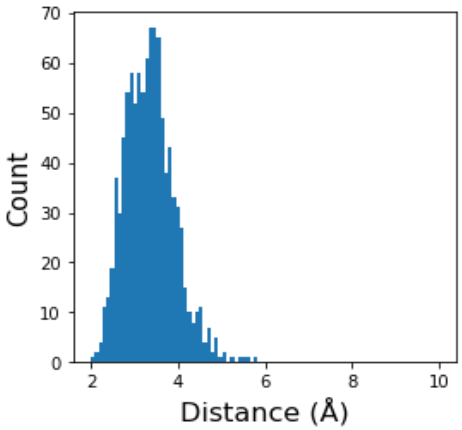
Average 4.4, min 3.7, max 5.4 Å

{1}ME (carbon) – 4Trp$^{\varepsilon 1}$



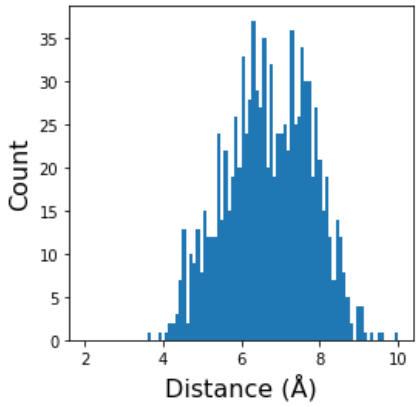
Average 3.5, min 2.4, max 5.3 Å

4Trp$^{\varepsilon 1}$ – {1}HA



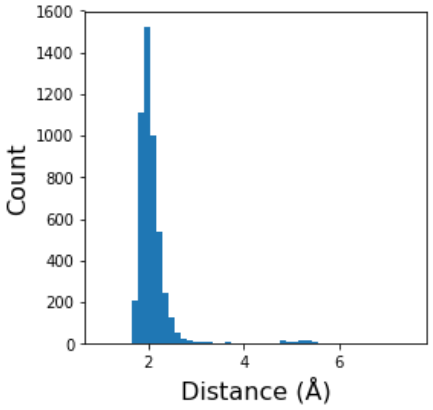
Average 3.4, min 2, max 5.8 Å

4Trp$^{\eta 2}$ – {1}HA



Average 6.7, min 3.7, max 10.4 Å

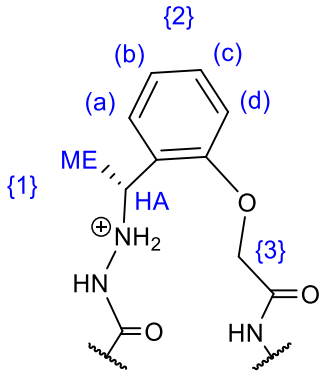
C=O (*i*) – NH(*i*+3)



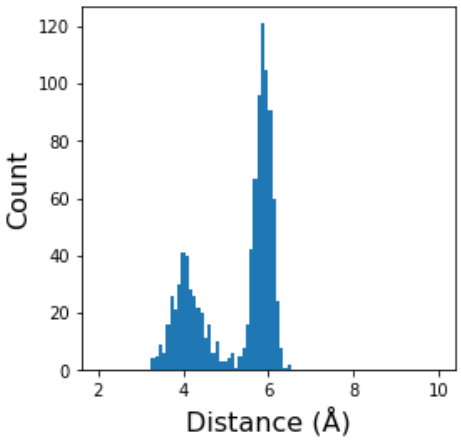
Average 2.11, min 1.63, max 6.36 Å

Figure 185. Minimum, maximum and average distances extracted from the MD simulation of (*S*)-15.

7.3.2.2 Distances extracted from the simulation of (*R*)-15

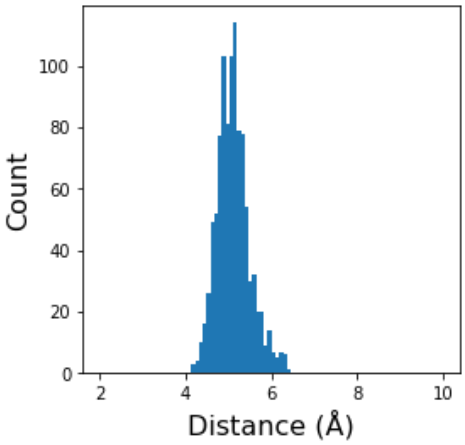


{1}ME (carbon) – {3}*pro*-R



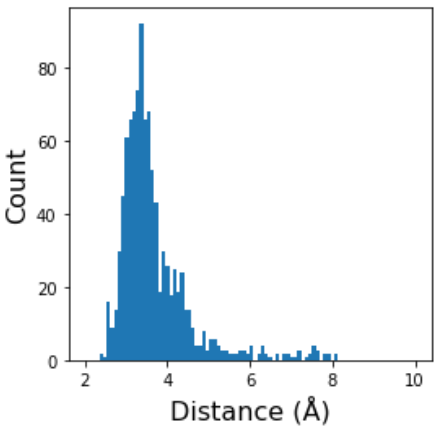
Average 5.3, min 3.2, max 6.5 Å

{1}ME (carbon) – {3}*pro*-R



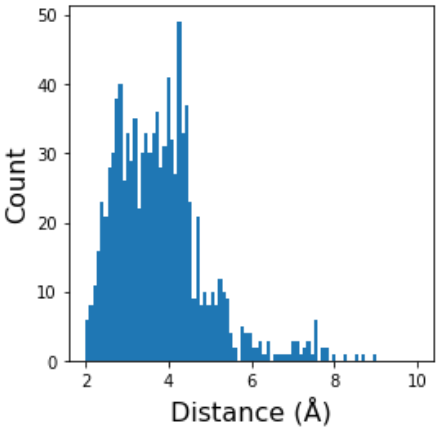
Average 5.1, min 4.2, max 6.4 Å

{1}ME (carbon) – 4Trp$^{\varepsilon 1}$



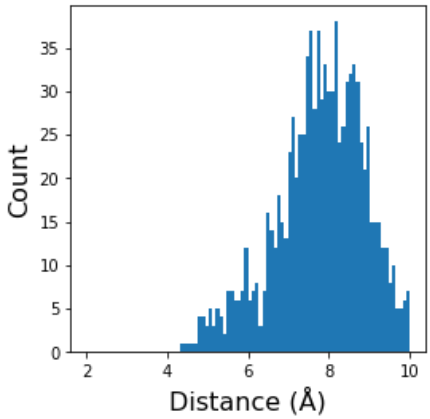
Average 3.7, min 2.4, max 8.1 Å

4Trp$^{\varepsilon 1}$ – {1}HA



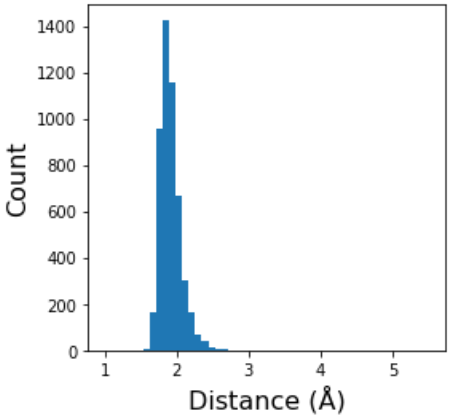
Average 3.8, min 1.9, max 9 Å

4Trp$^{\eta 2}$ – {1}HA



Average 7.9, min 4.4, max 11.2 Å

C=O (*i*) – NH(*i*+3)



Average 1.9, min 1.61, max 4.09 Å

Figure 186. Minimum, maximum and average distances extracted from the MD simulation of (*R*)-15.

7.4 SPR assays

All SPR experiments were performed by Dr Mads Gabrielsen (Prof Danny Huang's group, Beatson Institute) at 25°C on a Biacore T200 with a CM-5 chip (GE Healthcare). GST-tagged MDM2 1-109 were coupled to CM-5 chips as described previously.[358] **29** and **28** were serially diluted in running buffer containing 25 mM Tris-HCl, pH 7.6, 150 mM NaCl, 1 mM DTT and 0.005% (v/v) Tween-20. For experiments performed in the presence of control and mimic, both samples were diluted in running buffer with a final concentration of 10 mM of the peptide. This was used as stock for a 5-fold serial dilution ranging from 100 μM. Data reported are the difference in signal between GST-MDM2 1-109 variants and MDM2 1-109 alone.

7.5 General information

All reagents were purchased from commercial sources and used without further purification. Dry solvents were purified using a PureSolv 500 MD solvent purification system. Thin layer chromatography was performed on Merck alumina plates covered with silica gel 60 F_{254} and visualised using UV light and stained with potassium permanganate, *p*-anisaldehyde or ninhydrin solution. Normal phase column chromatography was performed on a Biotage Isolera One 3.0, using prepacked silica Biotage SNAP KP-Sil cartridges.

LC-MS analysis was performed on a Thermo Scientific Dionex Ultimate 3000 LC system coupled to a Thermo LCQ Fleet quadrupole mass spectrometer using positive mode electrospray ionisation (ESI+). The system was equipped with a Dr Maisch ReproSil Gold 120 C18, 110 Å, 3 µM, 150 x 4 mm column and a solvent system consisting of solution A (5% MeCN in H_2O + 0.1% TFA) and B (5% H_2O in MeCN + 0.1% TFA). Linear gradients were run from 0% to 100% buffer B over 10 or 20 min at a 1mL/min flow rate.

High resolution mass spectrometry (HRMS) was performed by the analytical service of the University of Glasgow, either on a Jeol M-Station JMS-700 High Resolution Mass Spectrometer using Electron ionisation or Chemical ionisation, or using ESI+ ionisation on a Bruker microTOF-Q II High Resolution Mass Spectrometer in positive mode. HRMS data are reported as mass to charge ratio (*m*/*z*) = observed/MW.

Nuclear magnetic resonance (NMR) spectra were recorded on a Bruker AVI 400MHz spectrometer (400 MHz for ^{1}H-NMR and 100 MHz for ^{13}C-NMR) or on a Bruker AVANCE 600 MHz spectrometer equipped with a TCI cryoprobe, at 298 K unless otherwise stated. NMR spectra for peptides were analysed using CCPN analysis.[359] Chemical shifts (δ_H) are reported in parts per million (ppm) relative to an internal standard (Me_4Si, TSP or DSS, 0 ppm) or the solvent peak ($CDCl_3$: 7.26 ppm, $(CD_3)_2SO$: 2.50 ppm, CD_3OD: 3.31 ppm, D_2O δ_H: 4.79 ppm). δ_C are reported in ppm relative to the signal of an internal standard (Me_4Si or TSP, 0 ppm) or the solvent peak ($CDCl_3$: 77.16 ppm, $(CD_3)_2SO$: 39.52 ppm, CD_3OD: 49.00 ppm). Proton and carbon chemical shifts were assigned using proton, carbon, Correlation Spectroscopy (COSY) and Heteronuclear Single Quantum Coherence (HSQC) experiments. Coupling constants (*J*) are reported in hertz (Hz). Splitting patterns are abbreviated as follows: singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m), broad (b), or a combination of these.

7.6 General information for peptide synthesis and analysis

Peptides were synthesised on a CEM Liberty Blue or on a Biotage Initiator+ Alstra microwave-assisted peptide synthesiser where specified, using the Fmoc/^{t}Bu protecting group strategy in a 0.1 mmol scale unless otherwise stated. All amino acids are of L-configuration unless otherwise stated. Compounds were frozen at -78°C and lyophilised on

a Christ Alpha 2-4 LO plus freeze dryer at -80°C and <0.5 mBar. Peptide content was analysed on a Thermo Scientific NanoDrop One UV-Vis spectrophotometer.

Peptides were purified on a Dionex P680 semi-preparative HPLC system equipped with P680 pumps and a UVD170U UV-Vis detector (monitoring at 214 nm and 280 nm), using either a Phenomenex, Gemini C18, 5 µm, 250 x 21.2 mm column at a flow rate of 6 or 8 mL/min or a Phenomenex, Luna C18(2), 5 µM, 100 Å, 150 x 10 mm column at a flow rate of 3 mL/min. Linear gradients were run using a binary solvent system consisting of solution A (H_2O + 0.1% TFA) and B (MeCN + 0.1% TFA).

Analytical RP-HPLC was performed on a Shimadzu reverse-phase HPLC system equipped with LC-20AT pumps, a SIL-20A autosampler and a SPD-20A UV-Vis detector (monitoring at 214 nm and 280 nm). The columns used were a Phenomenex Aeris 5 µm, peptide XB-C18, 150 x 4.6 mm or a Phenomenex Gemini-NX 5u C18 110Å 150 x 4:60 mm at a flow rate of 1 mL/min. Gradients were run using a binary solvent system consisting of solution A (5% MeCN in H_2O + 0.1% TFA) and B (5% H_2O in MeCN + 0.1% TFA) or using 100% solutions (A: H_2O + 0.1% TFA; B: MeCN + 0.1% TFA) when specified. Two or three gradients were used to determine the purity of each peptide: a gradient from 0% to 100% solution B over 20 min, a gradient from 0% to 50% solution B over 30 min or a gradient from 0% to 100% solution B over 50 min. Analytical HPLC data is reported as column retention time (RT) in minutes (min).

Circular dichroism measurements were performed with a JASCO J-810 circular dichroism spectropolarimeter fitted with a Peltier temperature controller, using a cuvette with a 0.1 or 0.2 cm pathlength.

7.7 General procedures for peptide synthesis

7.7.1 Microwave-assisted SPPS: CEM synthesiser

7.7.1.1 Resin swelling

The resin was swelled with 15 mL of DMF for 5 min at room temperature.

7.7.1.2 Fmoc deprotection

Fmoc deprotection was performed with 4 mL of a 20% morpholine in DMF solution, at rt and 0 W for 5 sec, 78 °C and 100 W for 30 sec, 88 °C and 70 W for 20 sec, and 90 °C and 25 W for 60 sec, followed by washing.

7.7.1.3 Coupling

Coupling of *N*-α-Fmoc protected amino acids (5 equiv., 0.2 M in DMF) was carried out using DIC (5 equiv., 0.5 M in DMF) and Oxyma Pure (5 equiv., 1 M in DMF). The reaction vessel was kept at 25 °C at 0 W for 5 sec, and then heated to 80 °C at 100 W for 30 sec, 86 °C at 70 W for 20 sec and 90 °C at 25 W for 120 sec. A wash was then performed.

Coupling of Fmoc-Cys(Trt)-OH was carried out without heating at 0 W for 120 sec and then at 50 °C at 50 W for 480 sec. The resin was then washed.

7.7.2 Microwave-assisted SPPS: Biotage synthesiser

Some peptides (where specified) were synthesised using this synthesiser and methods.

7.7.2.1 Resin swelling

The resin was swelled in 4.5 mL DMF (for a 0.1 mmol scale) at 70 °C for 20 min.

7.7.2.2 Fmoc deprotection

Fmoc deprotection was carried out in 20% morpholine in DMF, at rt for 30 s followed by 3 min at 70°C followed by washing with DMF.

7.7.2.3 Coupling

Coupling of *N*-α-Fmoc protected amino acids (4 equiv., 0.2 M in DMF) was carried out for 5 min at 75°C, using HCTU (4 equiv., 0.5 M in DMF) and DIPEA (8 equiv., 2 M in NMP), followed by washing with DMF.

Coupling of Fmoc-Cys(Trt)-OH (4 equiv.) was carried out in two cycles, first for 5 min at room temperature and then 5 min at 50 °C, followed by washing with DMF.

7.7.3 Manual resin swelling

Resins were swelled in 5 mL of DMF/DCM 1:1 for 30 min at room temperature.

7.7.4 Manual coupling

Coupling of *N*-α-Fmoc protected amino acids (5 equiv.) was carried out by dissolving the amino acid, DIC (4.5 equiv.) and Oxyma Pure (4.5 equiv.) in DMF (5 mL for a 0.1 mmol scale). DIPEA (6 equiv.) was then added. The solution was stirred for 5 min for preactivation and then it was added to the resin for reaction at room temperature (2.5 h).

7.7.5 Manual Fmoc deprotection

Deprotection was carried out for 15 min at room temperature with 5 mL (for a 0.1 mmol scale) of a 20% morpholine in DMF solution. The resin was then washed with DMF.

7.7.6 Acetyl capping

Capping of the *N*-terminus was carried out using acetic anhydride (50 equiv.) and pyridine (80 equiv.) for 30 min at room temperature followed by washing with DMF.

7.7.7 Cleavage test

A small portion of resin was transferred to a reaction vessel and 1 mL of a solution of TFA/H_2O/TIPS (95:2.5:2.5) was then added. The reaction was put on a rotary mixer for 45 min. After filtration, the solution was evaporated using a stream of nitrogen and the peptide was dissolved in H_2O/MeCN 1:1 and lyophilised.

7.7.8 Peptide cleavage

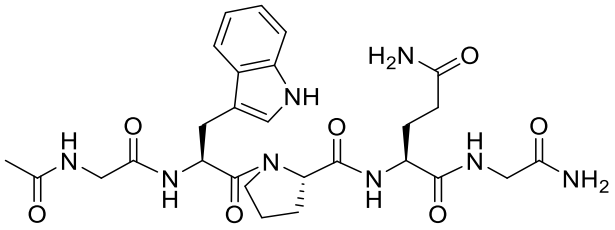
The resin-bound peptide was washed with DCM and the cleavage and removal of the protecting groups was then performed using a cleavage cocktail of 95% TFA, 2.5% H_2O and 2.5% TIPS (5 mL of solution for a 0.1 mmol scale). The reaction was left to progress for 2 h at room temperature on a mixer. After filtration, the cleavage cocktail was reduced in volume using a stream of nitrogen and the peptide was precipitated from solution with ice cold Et_2O (40 mL) and isolated via centrifugation at 3700 rpm for 5 min. The precipitate was washed with ice cold Et_2O, then centrifuged again and dissolved in H_2O/MeCN 1:1 and lyophilised.

7.7.9 TIPS-free peptide cleavage

The resin-bound peptide was washed with DCM and the cleavage and removal of the protecting groups was then performed using a cleavage cocktail of 95% TFA and 5% H_2O (5 mL of solution for a 0.1 mmol scale). The reaction was left to progress for 2 h at room temperature on a mixer. After filtration, the cleavage cocktail was reduced in volume using a stream of nitrogen and the peptide was precipitated from solution with ice cold Et_2O (40 mL) and isolated via centrifugation at 3700 rpm for 5 min. The precipitate was washed with ice cold Et_2O, then centrifuged again and dissolved in H_2O/MeCN 1:1 and lyophilised.

7.8 Chapter 2 Xaa-Pro-Yaa double mutant cycle peptides

Ac-GWPQG-NH_2, 1

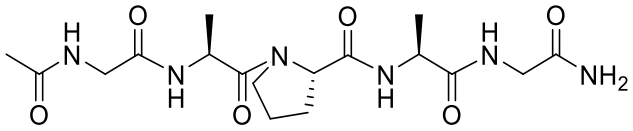


Scheme 69. Sequence of compound 1.

1 was synthesised on TentaGel S RAM resin (0.24 mmol/g loading) in a 0.1 mmol scale, using the general procedure for microwave-assisted SPPS on the Biotage synthesiser. After purification *via* RP-HPLC, the peptide was obtained as a white solid in a 7% yield.

HRMS (ESI+) calcd for $C_{27}H_{36}N_8O_7$ $[M+Na]^+$ 607.2599, found 607.2585. HPLC: 10.4 min 96.6% (20 min gradient), 17.9 min 94.7% (30 min gradient).

Ac-GAPAG-NH_2, 2

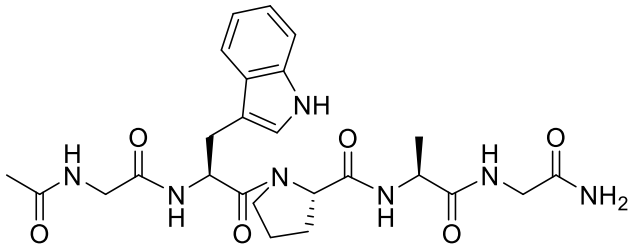


Scheme 70. Sequence of compound 2.

2 was synthesised on TentaGel S RAM resin (0.24 mmol/g loading) in a 0.1 mmol scale, using the general procedure for microwave-assisted SPPS on the Biotage synthesiser. After purification *via* RP-HPLC, the peptide was obtained as a white solid in a 16% yield.

HRMS (ESI+) calcd for $C_{17}H_{28}N_6O_6$ $[M+Na]^+$ 435.1963, found 435.1960. HPLC (run in 100% buffers): 8.8 min 85.5% (20 min gradient), 12.9 min 86.9% (30 min gradient).

Ac-GWPAG-NH_2, 3

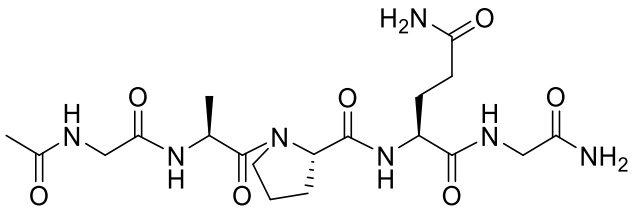


Scheme 71. Sequence of compound 3.

3 was synthesised on TentaGel S RAM resin (0.24 mmol/g loading) in a 0.1 mmol scale, using the general procedure for microwave-assisted SPPS on the Biotage synthesiser. After purification *via* RP-HPLC, the peptide was obtained as a white solid in a 9% yield.

HRMS (ESI+) calcd for $C_{25}H_{33}N_7O_6$ $[M+Na]^+$ 550.2385, found 550.2380. HPLC: 10.7 min 95.2% (20 min gradient), 18.7 min 97.0% (30 min gradient).

Ac-GAPQG-NH_2, 4

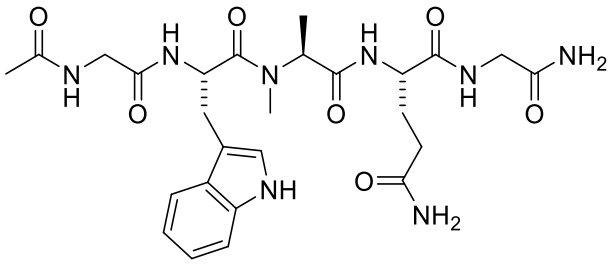


Scheme 72. Sequence of compound 4.

4 was synthesised on TentaGel S RAM resin (0.24 mmol/g loading) in a 0.1 mmol scale, using the general procedure for microwave-assisted SPPS on the Biotage synthesiser. After purification *via* RP-HPLC, the peptide was obtained as a white solid in a 2% yield.

HRMS (ESI+) calcd for $C_{43}H_{61}N_{11}O_9$ $[M+Na]^+$ 492.2177, found 492.2171. HPLC (run in 100% buffers): 8.6 min 92.3% (20 min gradient), 13.2 min 94.1% (30 min gradient).

Ac-GW(*N*-Me)AQG-NH_2, 5



Scheme 73. Sequence of compound 5.

5 was synthesised on TentaGel S Ram resin (0.24 mmol/g loading) in a 0.1 mmol scale. The first three amino acids in the sequence were coupled using the general procedure for microwave-assisted SPPS on the Biotage synthesiser.

o-NBS protection

Ortho-nitrobenzenesulfonyl chloride (*o*-NBS-Cl) (4 equiv.) was dissolved in NMP. Triethylamine (10 equiv.) was added to the solution and the mixture was added onto the resin. The reaction was left to progress for 15 min at room temperature.

N-methylation

After washing the resin with NMP, DBU (3 equiv.) was added to 1 mL NMP and the resin was treated with the solution for 3 min. Dimethylsulfate (DMS) (10 equiv.) was added to 1 mL NMP and the solution was added onto the resin. The reaction mixture was stirred for 2 min and the resin was then washed with NMP. This procedure was adapted from the literature.[219,249,250]

o-NBS deprotection

β-mercaptoethanol (BME) (10 equiv.) and DBU (5 equiv.) were added to 2 mL NMP. The resin was treated with the solution for 5 min and it was then washed with NMP.

Coupling of Trp on *N*-methylated Ala

Fmoc-Trp(Boc)-OH (5 equiv.) and HATU (4.5 equiv.) were dissolved in DMF. DIPEA (6 equiv.) was then added and the solution was stirred for 5 min. The resin was treated with the mixture for 3 h and washed with DMF.

The Fmoc group was removed following the general procedure for manual Fmoc deprotection. The solution was collected to allow for the evaluation of the Fmoc concentration. The remaining couplings, acetylation and cleavage were performed following the general manual procedures.

5 was obtained as a white solid in a 1% purified yield. HRMS (ESI+) calcd for $C_{26}H_{36}N_8O_7$ $[M+Na]^+$ 595.2599, found 595.2597. HPLC: 10.7 min 99.5% (20 min gradient), 18.6 min 99.4% (30 min gradient).

7.9 Chapter 2 GCxPCG disulfide peptides

Screening of conditions for disulfide formation

Disulfide formation conditions were screened on **6** as follows:

Air oxidation in NH_4HCO_3, 2.5 mM

Following the conditions from Calce *et al.*,[254] 7.6 µmol of **6** were dissolved in 3 mL of 0.1 M NH_4HCO_3 in a vial at a concentration of 2.5 mM. Air was then bubbled through the sample for 4 h. The reaction mixture was analysed by LCMS but no product or oligomer mass could be observed.

Air oxidation in NH_4HCO_3, 0.1 mM

Using conditions by Annis *et al.*,[360,361] 0.5 µmol of **6** were dissolved in NH_4HCO_3 (0.1 M, pH 8) at a concentration of 0.1 mM in a vial and air was bubbled through the sample for 43 h. A 92% yield for the formation of **8** was estimated by LCMS.

NCS-mediated oxidation

Following the conditions described by Albericio *et al.*,[255,362] a solution of *N*-chlorosuccinimide in H_2O/MeCN 1:1 was prepared and then added onto 0.5 µmol of **6** (2 equiv. NCS, peptide concentration 0.1 mM) in a vial. The reaction proceeded for 15 min and it was then lyophilised. An estimated yield of 83% for the formation of **8** was found by LCMS.

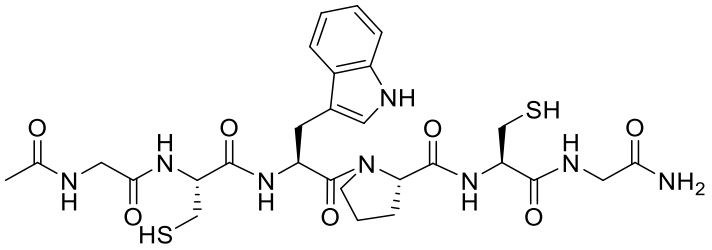
DMSO-mediated oxidation

In a vial, 0.5 µmol of **6** were dissolved in NH_4HCO_3 0.1 M (0.1 mM peptide concentration, pH 8) and 15% in volume of DMSO was added.[363] The reaction was left to stir for 6 h and it was analysed by LCMS. A yield of 88% for the formation of **8** was estimated by LCMS.

DPDS-mediated oxidation

In a vial, 0.2 µmol of **6** were dissolved in NH_4HCO_3 0.1 M (0.1 mM peptide concentration, pH 8) and 3 equivalents of 2,2'-dipyridyldisulfide were added from a 1 mM solution in MeOH.[257] The reaction was left for 1 h and then stopped by acidification with TFA. LCMS analysis showed an estimated yield of 95% for the formation of **8**.

Ac-GCWPCG-NH_2, 6

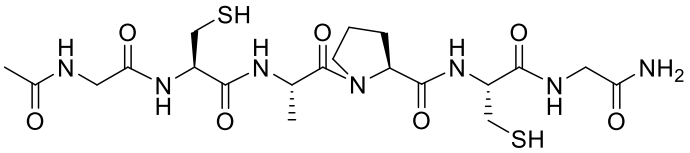


Scheme 74. Sequence of compound 6.

6 was synthesised on TentaGel S RAM resin (0.24 mmol/g loading) in a 0.1 mmol scale, using the general procedure for microwave-assisted SPPS on the Biotage synthesiser. The peptide was obtained as a white solid in a 34% yield after purification *via* RP-HPLC.

HRMS (ESI+) calcd for $C_{28}H_{38}N_8O_7S_2$ $[M+Na]^+$ 685.2197, found 685.2194. HPLC: 11.9 min 98.2% (20 min gradient), 22.2 min 97.7% (30 min gradient).

Ac-GCAPCG-NH_2, 7

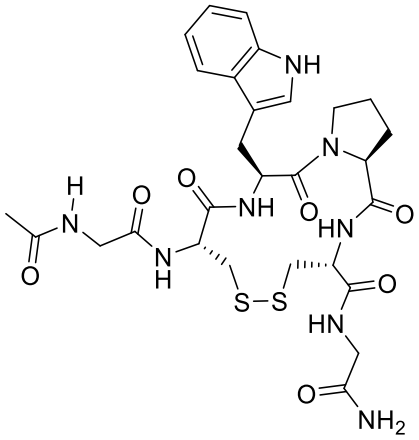


Scheme 75. Sequence of compound 7.

7 was synthesised on TentaGel S RAM resin (0.24 mmol/g loading) in a 0.1 mmol scale, using the general procedure for manual SPPS. The peptide was obtained as a white solid in a 31% yield after purification *via* RP-HPLC.

HRMS (ESI+) calcd for $C_{20}H_{33}N_7O_7S_2$ $[M+Na]^+$ 570.1775, found 570.1770. HPLC: 9 min 99.6% (20 min gradient), 13.4 min 95.3% (30 min gradient).

Disulfide-bridged Ac-GCWPCG-NH_2, 8

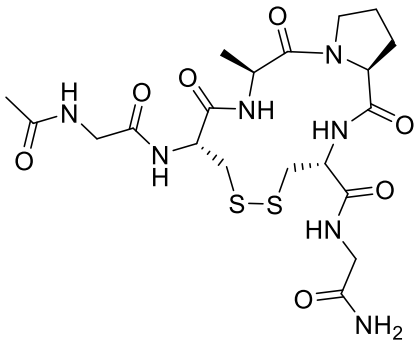


Scheme 76. Sequence of compound 8.

In a vial, 0.2 µmol of **6** were dissolved in aqueous NH_4HCO_3 (0.1 M, pH 8) at a 0.1 mM concentration and 3 equivalents of 2,2'-dipyridyldisulfide were added from a 1 mM solution in MeOH.[257] The reaction was stirred for 1 h and then stopped by acidification with TFA. After purification *via* RP-HPLC, **8** was obtained as a white solid in a 56% yield.

HRMS (ESI+) calcd for $C_{28}H_{36}N_8O_7S_2$ $[M+Na]^+$ 683.2041, found 683.2036. HPLC: 11.3 min 90.9% (20 min gradient), 20.4 min 88.9% (30 min gradient).

Disulfide-bridged Ac-GCAPCG-NH_2, 9



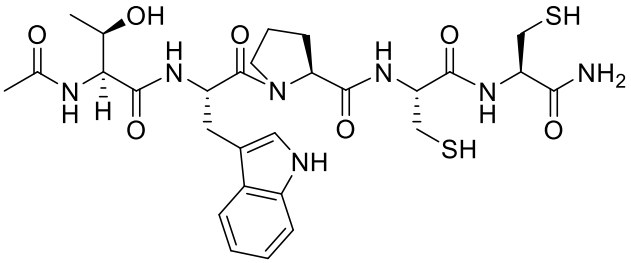
Scheme 77. Sequence of compound 9.

9 was synthesised following the same DPDS-mediated procedure as for **8**. After purification *via* RP-HPLC, the peptide was obtained as a white solid in a 50% yield.

HRMS (ESI+) calcd for $C_{20}H_{31}N_7O_7S_2$ $[M+Na]^+$ 568.1619, found 568.1601. HPLC: 9.1 min 97.9% (20 min gradient), 14.2 min 98.2% (30 min gradient).

7.10 Chapter 2 TWxCC vicinal disulfide peptides

Ac-TWPCC-NH_2, 10

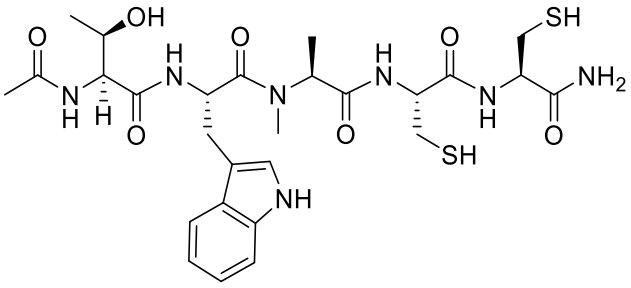


Scheme 78. Sequence of compound 10.

10 was synthesised on Fmoc-Rink Amide AM resin (0.74 mmol/g loading) in a 0.1 mmol scale, using the general procedure for microwave-assisted SPPS. After purification *via* RP-HPLC, the peptide was obtained as a white solid in a 33% yield.

HRMS (ESI+) calcd for $C_{28}H_{39}N_7O_7S_2$ $[M+Na]^+$ 672.2245, found 672.2238. HPLC: 12.2 min 97.5% (20 min gradient), 23 min 98.4% (30 min gradient).

Ac-TW(*N*-Me)ACC-NH_2, 11

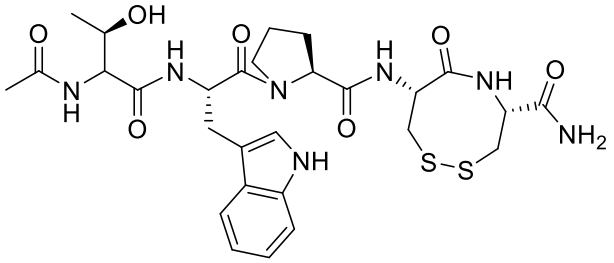


Scheme 79. Sequence of compound 11.

11 was synthesised on Fmoc-Rink Amide AM resin (0.74 mmol/g loading) in a 0.25 mmol scale, using the general procedure for microwave-assisted SPPS. The Trp following the *N*-methylated residue was coupled twice. After purification *via* RP-HPLC, the peptide was obtained as a white solid in a 10% yield.

HRMS (ESI+) calcd for $C_{27}H_{39}N_7O_7S_2$ $[M+Na]^+$ 660.2245, found 660.2255. HPLC: 12.4 min 98.9% (20 min gradient), 23.2 min 90.4% (30 min gradient).

Disulfide-bridged Ac-TWPCC-NH_2, 12

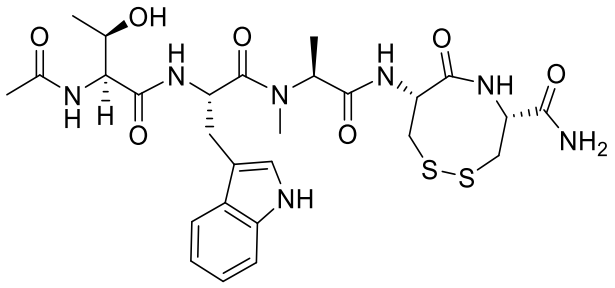


Scheme 80. Sequence of compound 12.

In a microwave vial, **10** (4 μmol, 1 equiv.) was dissolved in NH_4HCO_3 buffer (20 mM, pH 7) to a concentration of 0.2 mM. 2,2'-dipyridyldisulfide (3 equiv) was dissolved in MeOH (12 mL) and then added onto the peptide solution. The reaction was stirred at 50 °C for 1 h. After that time, the reaction was quenched by acidification with TFA until the pH was acidic. After purification *via* RP-HPLC, **12** was obtained in a 43% yield.

HRMS (ESI+) calcd for $C_{28}H_{37}N_7O_7S_2$ $[M+Na]^+$ 670.2088, found 670.2069. HPLC: 12.1 min 99% (20 min gradient), 22.6 min 98.5% (30 min gradient).

Disulfide-bridged Ac-TW(*N*-Me)ACC-NH_2, 13



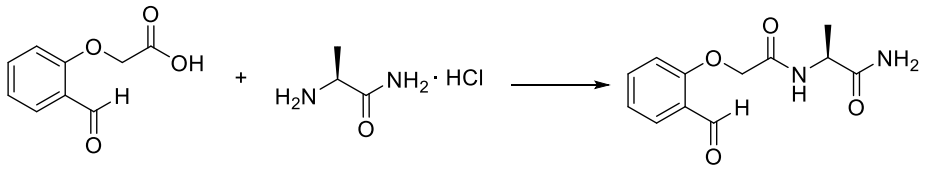
Scheme 81. Sequence of compound 13.

13 was synthesised from **11** in a 1 μmol scale, following the same procedure as for **12**. This afforded **13** in a 22% yield after purification.

HRMS (ESI+) calcd for $C_{27}H_{37}N_7O_7S_2$ $[M+Na]^+$ 658.2088, found 658.2074. HPLC: 12 min 98.2% (20 min gradient), 22.4 min 99.5% (30 min gradient).

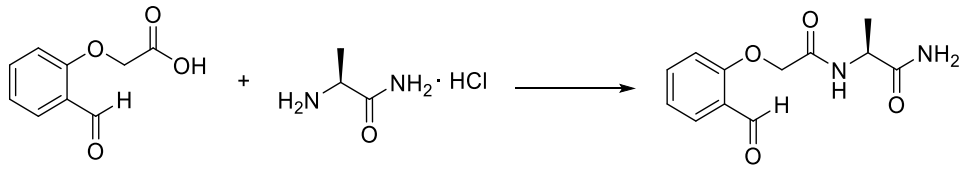
7.11 Chapter 3 tripeptide systems

Solution synthesis test of (2*S*)-2-[2-(2-formylphenoxy)acetamido] propanamide, 22



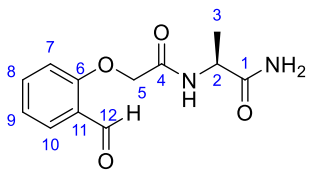
2-Formylphenoxyacetic acid (1 equiv., 1 mmol) and Oxyma Pure (1.5 equiv.) were dissolved in DMF (10 mL). H-Ala-NH_2·HCl (1.5 equiv.) was then added and the reaction mixture was cooled to 0 °C. Once cooled, DIPEA (2.5 equiv.) and EDCI·HCl (1.5 equiv.) were added and the reaction was left to stir overnight at room temperature. The reaction mixture was concentrated under vacuum and then DCM was added (10 mL) – the organic layer was washed with aqueous LiCl 1 M (1 x 10 mL) and aqueous NaOH 1 M (3 x 10 mL) and dried over $MgSO_4$. The mixture was then filtered and the solvent was evaporated *in vacuo*. The crude was purified through automated flash chromatography using a 0 to 15% MeOH/DCM gradient to furnish **22** as a white solid (4 mg, 2% yield). For the characterisation data, see the final procedure.

Base-free solution synthesis test of (2*S*)-2-[2-(2-formylphenoxy) acetamido]propanamide, 22



2-Formylphenoxyacetic acid (1 equiv., 1 mmol), Oxyma Pure (1.2 equiv.) and Ala-NH_2·HCl (1 equiv.) were dissolved in DMF. The mixture was cooled to 0 °C and then EDCI·HCl (1.1 equiv.) was added. The ice bath was then removed and the reaction was stirred at rt overnight. The reaction mixture was concentrated under vacuum and re-dissolved in DCM (10 mL) – the organic layer was washed with aqueous LiCl 1 M (1 x 10 mL) and sat. $NaHCO_3$ (3 x 10 mL) and dried over $MgSO_4$. The mixture was then filtered and the solvent was evaporated *in vacuo*. The crude was purified through automated flash chromatography using 10% MeOH/DCM to furnish **22** as an orange solid (15.2 mg, 6% yield). For the characterisation data, see the final procedure.

Final on-resin synthesis of (2*S*)-2-[2-(2-formylphenoxy)acetamido] propanamide, 22

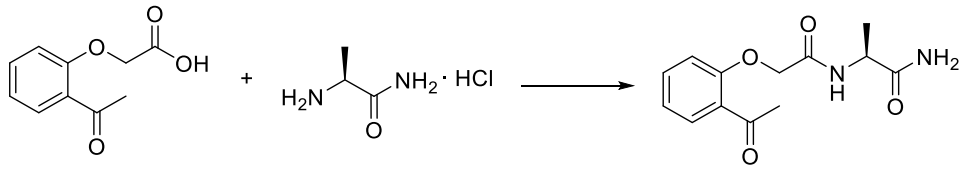


0.3 mmol of Fmoc-Rink Amide AM Resin (0.74 mmol/g loading) were swelled following the general procedure. Following the general procedures for manual SPPS, the resin was deprotected and Fmoc-Ala-OH·H_2O was coupled, followed by Fmoc deprotection. 2-Formylphenoxyacetic acid (5 equiv.) and Oxyma Pure (4.5 equiv.) were dissolved in DMF and then DIPEA (6 equiv.) and DIC (4.5 equiv.) were added. The solution was preactivated for 5 min and it was then added for reaction for 2.5 h at rt. Cleavage from the resin was performed following the general procedure, using only water as the scavenger. Purification was performed through RP-HPLC using a 20 to 60% B gradient. **22** was obtained as a white solid (34.1 mg, 45% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 10.16 (s, 1H, 12-H), 8.17 (d, *J* = 6.0 Hz, 1H, NH), 7.79 (dd, *J* = 7.5, 1.8 Hz, 1H, Ar-H), 7.61 (ddd, *J* = 8.3, 7.5, 1.8 Hz, 1H, Ar-H), 7.21 (td, *J* = 7.5, 0.8 Hz, 1H, Ar-H), 6.94 (br d, *J* = 8.3, 0.8 Hz, 1H, Ar-H), 6.27 (s, 1H, N*H*^{A}H^B), 5.42 (s, 1H,

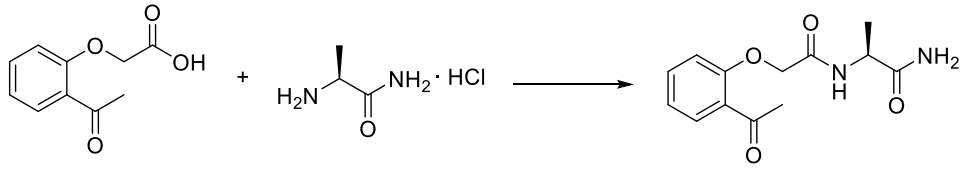
$NH^A H^B$), 4.62 (s, 2H, 5-H), 4.61 – 4.54 (m, 1H, 2-H), 1.56 (d, J = 7.1 Hz, 3H, 3-H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 190.43 (CH), 173.77 (C=O), 167.91 (C=O), 157.58 (C), 136.12 (Ar-CH), 134.56 (Ar-CH), 125.12 (C), 122.26 (Ar-CH), 113.24 (Ar-CH), 67.58 (CH_2), 48.51 (CH), 17.33 (CH_3). HRMS (ESI^+) calcd for $C_{12}H_{14}N_2O_4$ $[M+Na]^+$ 273.0846, found 273.0846.

Solution synthesis test of (2*S*)-2-[2-(2-acetylphenoxy)acetamido] propanamide, 21



In a 25 mL round bottom flask, 2-acetylphenoxyacetic acid (1 equiv., 1 mmol), Oxyma Pure (1.5 equiv.) and DMF (10 mL) were added. H-Ala-NH_2·HCl (1.5 equiv.) was then added and the reaction mixture was cooled to 0 °C. After cooling, DIPEA (2.5 equiv.) and EDCI·HCl (1.5 equiv.) were added and the reaction was stirred overnight at room temperature. The reaction mixture was concentrated under vacuum and then DCM was added (10 mL) – the organic layer was washed with aqueous LiCl 1 M (1 x 10 mL) and aqueous NaOH 1 M (3 x 10 mL) and dried over $MgSO_4$. The mixture was then filtered and the solvent was evaporated *in vacuo*. The crude was purified through automated flash chromatography using a 0 to 15% MeOH/DCM gradient to furnish **21** as a white solid (79.5 mg, 30% yield). For the characterisation data, see the final procedure.

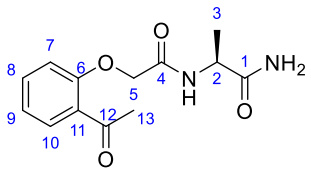
Base-free solution synthesis test of (2*S*)-2-[2-(2-acetylphenoxy) acetamido]propanamide, 21



2-Acetylphenoxyacetic acid (1 equiv., 1 mmol), Oxyma Pure (1.2 equiv.) and Ala-NH_2·HCl (1 equiv.) were dissolved in DMF. The mixture was cooled to 0 °C and then EDCI·HCl (1.1 equiv.) was added. The ice bath was then removed and the reaction was stirred at rt overnight. The reaction mixture was concentrated under vacuum and re-dissolved in DCM (10 mL) – the organic layer was washed with aqueous LiCl 1 M (1 x 10 mL) and sat. $NaHCO_3$ (3 x 10 mL) and dried over $MgSO_4$. The mixture was then filtered and the solvent

was evaporated *in vacuo*. The crude was purified through automated flash chromatography using 10% MeOH/DCM to furnish **21** as a yellow solid (82.5 mg, 31% yield). For the characterisation data, see the final procedure.

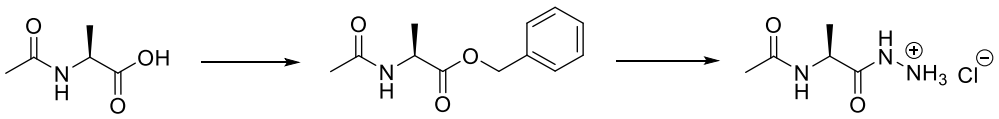
Final on-resin synthesis of (2*S*)-2-[2-(2-acetylphenoxy)acetamido] propanamide, 21



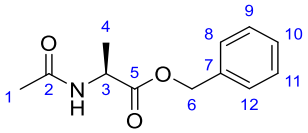
0.1 mmol of Fmoc-Rink Amide AM Resin (0.74 mmol/g loading) were swelled following the general procedure. Following the general procedures for manual SPPS, the resin was deprotected and Fmoc-Ala-OH·H_2O was coupled, followed by Fmoc deprotection. 2-Acetylphenoxyacetic acid (5 equiv.) and Oxyma Pure (4.5 equiv.) were dissolved in DMF and then DIPEA (6 equiv.) and DIC (4.5 equiv.) were added. The solution was preactivated for 5 min and it was then added for reaction for 2.5 h at rt. Cleavage from the resin was performed following the general procedure, using only water as the scavenger. Purification was performed through RP-HPLC using a 20 to 60% B gradient. **21** was obtained as a white solid (17.8 mg, 67% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.38 (d, *J* = 7.5 Hz, 1H, NH), 7.72 (dd, *J* = 7.5, 1.7 Hz, 1H, Ar-H), 7.45 (ddd, *J* = 8.3, 7.5, 1.7 Hz, 1H, Ar-H), 7.04 (td, *J* = 7.5, 1.0 Hz, 1H, Ar-H), 6.87 (dd, *J* = 8.3, 1.0 Hz, 1H, Ar-H), 6.29 (s, 1H, NH^AH^B), 5.34 (s, 1H, NH^AH^B), 4.54 (s, 2H, 5-H), 4.50 (q, *J* = 7.3 Hz, 1H, 2-H), 2.56 (s, 3H, 13-H), 1.47 (d, *J* = 7.1 Hz, 3H, 3-H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 198.90 (C=O), 173.99 (C=O), 168.43 (C=O), 156.28 (Ar-C), 134.35 (Ar-CH), 131.43 (Ar-CH), 127.23 (Ar-C), 121.79 (Ar-CH), 113.85 (Ar-CH), 67.87 (CH_2), 48.56 (CH), 29.73 (CH_3), 17.17 (CH_3). HRMS (ESI^+) calcd for $C_{13}H_{16}N_2O_4$ $[M+Na]^+$ 287.1002, found 287.1006.

Test synthetic route A for the synthesis of *N*-[(1*S*)-1-(hydrazinecarbonyl)ethyl]acetamide, 23



Benzyl (2*S*)-2-acetamidopropanoate, 24

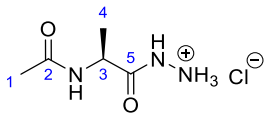


N-acetyl-L-alanine (1 equiv., 2.3 mmol), triethylamine (1.2 equiv.) and benzyl bromide (1.2 equiv.) were dissolved in acetonitrile (25 mL). The reaction mixture was left to stir at room temperature overnight. The crude was then concentrated under vacuum and chloroform/water 1:1 were added (30 mL). The aqueous layer was then washed with chloroform (2 x 15 mL) and the combined organic layers were washed with water (2 x 15 mL) and a saturated NaCl solution (1 x 15 mL). The organic layers were then dried over $MgSO_4$, filtered and the solvent was removed under vacuum to furnish benzyl acetyl-L-alaninate as a clear oil (352.4 mg, 81% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.32 – 7.18 (m, 5H, Ar-H), 6.34 (d, *J* = 7.6 Hz, 1H, NH), 5.08 (d, *J* = 3.7 Hz, 2H, 6-H), 4.59 – 4.52 (m, 1H, 3-H), 1.91 (s, 3H, 1-H), 1.31 (d, *J* = 7.3 Hz, 3H, 4-H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 173.07 (C=O), 169.75 (C=O), 135.38 (Ar-C), 128.61 (2 x Ar-CH), 128.42 (Ar-CH), 128.08 (2 x Ar-CH), 67.07 (-CH_2-), 48.14 (CH), 23.02 (CH_3), 18.34 (CH_3). HRMS (ESI+) calcd for $C_{12}H_{15}NO_3$ $[M+Na]^+$ 244.0944, found 244.0944.

The spectroscopic data was in good agreement with the literature.[364]

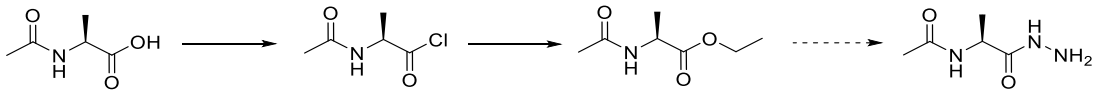
[(2*S*)-2-acetamidopropanamido]azanium chloride, 25



24 (1 equiv., 0.18 mmol) and EtOH (2 mL) were added to a sealed microwave vial. Hydrazine hydrate (10 equiv.) was then added and the mixture was refluxed at 80 °C for 6 h. The reaction crude was concentrated under vacuum. In order to make the HCl salt, acetyl chloride (10 equiv.) was added dropwise onto MeOH. The mixture was stirred for 10 min and it was then added onto the reaction crude. The reaction was stirred for 10 min at rt and then the mixture was dried *in vacuo*. The colourless oil obtained was triturated with cold MeCN and filtrated using a Buchner funnel and frit. [(2*S*)-2-acetamidopropanamido]azanium chloride was obtained as a white solid. The yield was found to be 76% (32.6 mg), but this could be lower as residual hydrazine could be co-precipitating with the desired product. For this reason, synthetic route A was abandoned and routes B and C were to be explored instead.

^{1}H NMR (400 MHz, DMSO-d_6) δ 11.03 (s, 1H, N*H*-NH$_3$), 8.29 (d, *J* = 6.5 Hz, 1H, Ac-N*H*), 4.32 – 4.25 (m, 2H, 3-H), 1.84 (s, 3H, 1-H), 1.24 (d, *J* = 7.2 Hz, 3H, 4-H). ^{13}C NMR data could not be obtained due to lack of material. HRMS (ESI+) calcd for $C_5H_{11}N_3O_2$ $[M+Na]^+$ 168.0743, found 168.0744.

Test synthetic route B for the synthesis of *N*-[(1*S*)-1-(hydrazinecarbonyl)ethyl]acetamide, 23



Conditions for the first step at room temperature

Acetyl alanine (2.3 mmol, 1 equiv.) was dissolved in dry DMF (25 mL) under a nitrogen atmosphere. The solution was cooled to 0 °C and then $SOCl_2$ (1.2 equiv.) was added dropwise. The ice bath was then removed and the reaction mixture was stirred at rt overnight. After that time, the reaction crude was concentrated under vacuum and then re-dissolved in EtOH (25 mL). The mixture was refluxed at 80 °C for 2 h. After concentration *in vacuo*, the crude was dissolved in DCM and washed with water (1 x 10 mL), aqueous LiCl 1 M (1 x 10

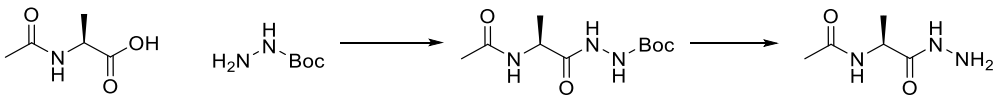
mL) and sat. NaCl (1 x 10 mL) and then dried over $MgSO_4$. After concentration under vacuum a brown oil was obtained.

Conditions for the first step under reflux

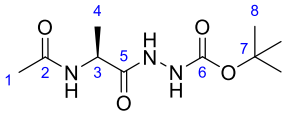
Acetyl alanine (2.3 mmol, 1 equiv.) and $SOCl_2$ (1 equiv.) were mixed and the reaction was refluxed at 80 °C for 2 h. After concentration *in vacuo*, the crude was re-dissolved in EtOH (10 mL) and the mixture was refluxed at 80 °C for 2 h.

Using these conditions no product formation was observed, and therefore this synthetic strategy was abandoned. Strategy C was explored instead.

Final strategy C for the synthesis of *N*-[(1*S*)-1-(hydrazinecarbonyl) ethyl]acetamide, 23



***N*-[(1*S*)-1-{*N*'-[(*tert*-butoxy)carbonyl]hydrazinecarbonyl} ethyl]acetamide, 26**

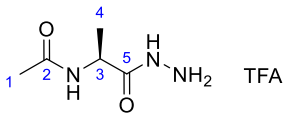


N-acetyl-L-alanine (1 equiv., 5 mmol) and *tert*-butyl carbazate (1.7 equiv.) were dissolved in 10% dry DMF/dry DCM (20 mL). Oxyma Pure (1.4 equiv.) and EDCI HCl (1.1 equiv.) were then added and the reaction was stirred at rt for 20 h.[275] After concentration *in vacuo*, the crude was purified through column chromatography eluting with 92:6:2 DCM/MeOH/acetone, which furnished **26** as a yellow oil (0.66 g, 54% yield).

^{1}H NMR (400 MHz, CD_3OD) δ 4.33 – 4.45 (m, 1H, 3-H), 1.97 (s, 3H, 1-H), 1.46 (s, 9H, 8-H), 1.37 (d, *J* = 7.2 Hz, 3H, 4-H). ^{13}C NMR (101 MHz, CD_3OD) δ 173.33 (C=O), 171.61 (C=O), 156.16 (C=O), 80.43 (C), 47.47 (CH), 27.13 (CH_3), 21.03 (CH_3), 16.68 (CH_3). HRMS (ESI+) calcd for $C_{10}H_{19}N_3O_4$ $[M+Na]^+$ 268.1268, found 268.1267.

The spectroscopic data was in good agreement with the literature.[275]

***N*-[(1*S*)-1-(hydrazinecarbonyl)ethyl]acetamide, 23**

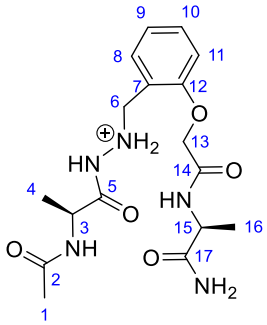


26 (1 equiv., 1 mmol) was dissolved in DCM (16 mL) and TFA (20% volume) was then added. The reaction was stirred for 1 h at rt and the crude was then concentrated under vacuum.[275] The oil obtained was re-dissolved in Et_2O and concentrated under vacuum again to afford **23** as a white solid (0.27 g, quantitative yield).

^{1}H NMR (400 MHz, CD_3OD) δ 4.33 (q, *J* = 7.2 Hz, 1H, 3-H), 1.98 (s, 3H, 1-H), 1.38 (d, *J* = 7.2 Hz, 3H, 4-H). ^{13}C NMR (101 MHz, CD_3OD) δ 172.58 (C=O), 172.01 (C=O), 47.80 (CH), 20.85 (CH_3), 16.14 (CH_3). HRMS (ESI+) calcd for $C_5H_{11}N_3O_2$ $[M+Na]^+$ 168.0743, found 168.0746.

The spectroscopic data was in good agreement with the literature.[275]

Ac-Ala-(U-BTM)-Ala-NH_2, 19

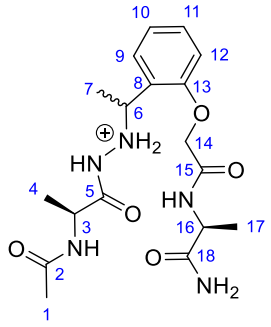


22 (1 equiv., 0.06 mmol) and **23** (1.2 equiv.) were dissolved in MeOH/AcOH 1:1 (40 mL). Sodium cyanoborohydride (10 equiv.) was then added and the reaction was stirred for 30 min at rt. Purification was performed through RP-HPLC using a 20 to 50% B gradient, which afforded **19** as a white solid (11 mg, 50% yield).

^{1}H NMR (400 MHz, CD_3OD) δ 7.41 (ddd, *J* = 8.3, 7.7, 1.7 Hz, 1H, Ar-H), 7.34 (dd, *J* = 7.7, 1.7 Hz, 1H, Ar-H), 7.06 – 7.02 (m, 2H, Ar-H), 4.78 (d, *J* = 1.2 Hz, 2H, 6-H), 4.48 (q, *J* = 7.2 Hz, 1H, 15-H), 4.31 (s, 2H, 13-H), 4.25 (q, *J* = 7.2 Hz, 1H, 3-H), 1.96 (s, 3H, 1-H), 1.42 (d, *J* = 7.2 Hz, 3H, 16-H), 1.26 (d, *J* = 7.2 Hz, 3H, 4-H). ^{13}C NMR (101 MHz, CD_3OD) δ 177.39 (C=O), 173.47 (C=O), 173.34 (C=O), 171.21 (C=O), 158.10 (Ar-C), 133.74 (Ar-CH), 132.28 (Ar-CH), 123.07 (Ar-CH), 122.21 (Ar-C), 113.38 (Ar-CH), 68.01 (CH_2), 52.15

(CH_2), 50.02 (CH), 49.21 (CH), 22.30 (CH_3), 18.33 (CH_3), 17.63 (CH_3). HRMS (ESI+) calcd for $C_{17}H_{25}N_5O_5$ $[M+Na]^+$ 402.1748, found 402.1753.

Ac-Ala-((*S*)-BTM)-Ala-NH_2 and Ac-Ala-((*R*)-BTM)-Ala-NH_2, 20a and 20b



21 (1 equiv., 0.06 mmol) and **23** (1.2 equiv.) were dissolved in MeOH/AcOH 1:1 (40 mL). Sodium cyanoborohydride (10 equiv.) was then added and the reaction was stirred for 30 min at rt. Purification was performed through RP-HPLC using a 20 to 50% B gradient, which afforded **20a** and **20b** as white solids in a proportion of 29% **20a** to 71% **20b** (total 13.3 mg, 56% overall yield).

20a

^{1}H NMR (400 MHz, CD_3OD) δ 7.38 – 7.31 (m, 2H, Ar-H), 7.04 (d, *J* = 7.4 Hz, 1H, Ar-H), 7.02 – 6.97 (m, 1H, Ar-H), 4.80 (s, 2H, 14-H), 4.77 – 4.74 (m, 1H, 6-H), 4.47 (q, *J* = 7.2 Hz, 1H, 16-H), 4.21 (q, *J* = 7.1 Hz, 1H, 3-H), 1.57 (d, *J* = 6.9 Hz, 3H, 7-H), 1.43 (d, *J* = 7.2 Hz, 3H, 17-H), 1.30 (s, 3H, 1-H), 1.13 (d, *J* = 7.1 Hz, 3H, 4-H). ^{13}C NMR (101 MHz, CD_3OD) δ 175.98 (2 x C=O), 171.72 (2 x C=O), 155.99 (Ar-C), 129.82 (Ar-CH), 129.42 (Ar-CH), 121.73 (Ar-C), 121.66 (Ar-CH), 112.11 (Ar-CH), 66.60 (CH_2), 56.65 (CH), 48.65 (CH), 47.81 (CH), 22.80 (CH_3), 16.98 (CH_3), 16.31 (CH_3), 15.61 (CH_3). HRMS (ESI+) calcd for $C_{18}H_{27}N_5O_5$ $[M+Na]^+$ 416.1904, found 416.1897.

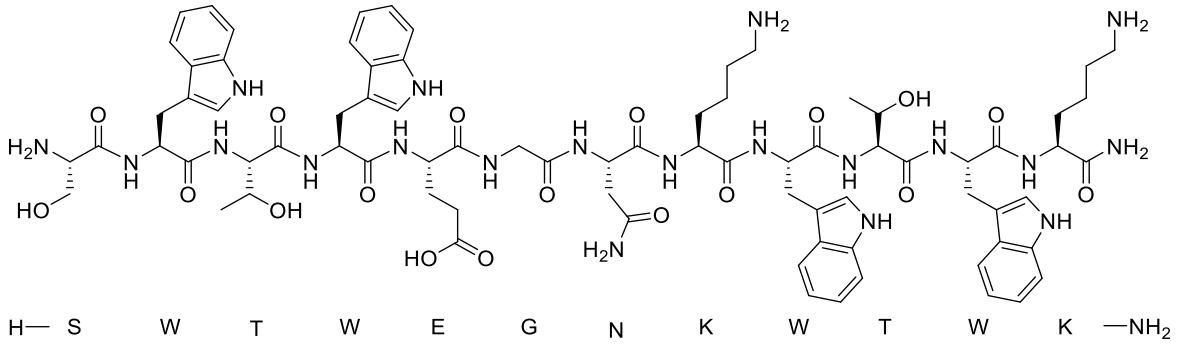
20b

^{1}H NMR (400 MHz, CD_3OD) δ 7.37 (m, 2H, Ar-H), 7.08 – 6.99 (m, 2H, Ar-H), 4.81 (s, 2H, 14-H), 4.77 – 4.71 (m, 1H, 6-H), 4.48 (q, *J* = 7.2 Hz, 1H, 16-H), 4.23 (q, *J* = 7.2 Hz, 1H, 3-H), 1.94 (s, 3H, 1-H), 1.59 (d, *J* = 6.9 Hz, 3H, 7-H), 1.43 (d, *J* = 7.2 Hz, 3H, 17-H), 1.21 (d, *J* = 7.2 Hz, 3H, 4-H). ^{13}C NMR (101 MHz, CD_3OD) δ 175.95 (C=O), 172.00 (C=O), 171.87 (C=O), 169.75 (C=O), 156.04 (Ar-C), 130.19 (Ar-CH), 129.30 (Ar-CH), 125.53 (Ar-C),

121.80 (Ar-CH), 112.29 (Ar-CH), 66.65 (CH), 56.74 (CH_2), 48.65 (CH), 47.89 (CH), 20.93 (CH_3), 17.01 (CH_3), 16.28 (CH_3), 15.25 (CH_3). HRMS (ESI+) calcd for $C_{18}H_{27}N_5O_5$ $[M+Na]^+$ 416.1904, found 416.1910.

7.12 Chapter 3 peptides

TrpZip1

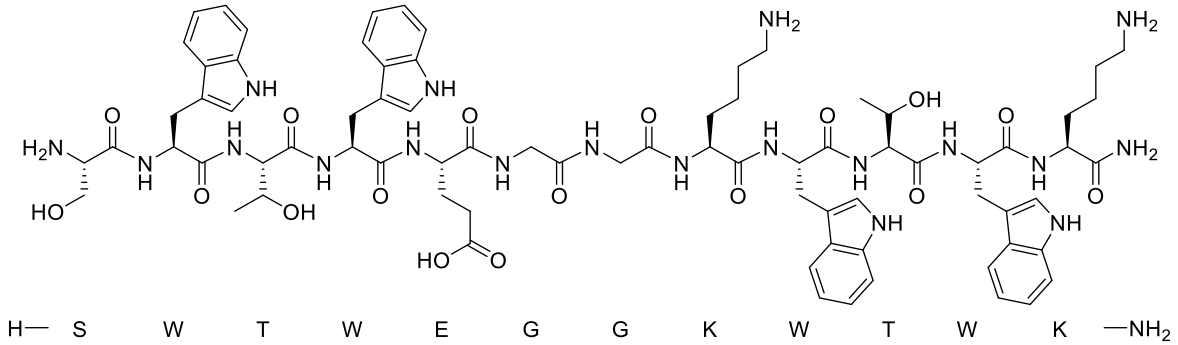


Scheme 82. Sequence of TrpZip1.

TrpZip1 was synthesised following the general procedure for microwave-assisted SPPS, using TentaGel S RAM resin (0.24 mmol/g loading) in a 0.1 mmol scale. After purification *via* RP-HPLC, TrpZip1 was obtained in a 14% yield.

HRMS (ESI+) calcd for $C_{78}H_{104}N_{20}O_{18}$ $[M+2H]^{2+}$ 804.3913, found 804.3883. HPLC: 12 min 100% (20 min gradient), 23.7 min 99.4% (30 min gradient).

TrpZipGG

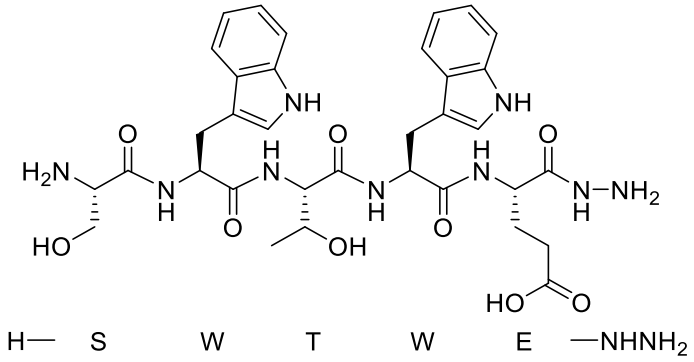


Scheme 83. Sequence of TrpZipGG.

TrpZipGG was synthesised following the general procedure for microwave-assisted SPPS, using TentaGel S RAM resin (0.24 mmol/g loading) in a 0.1 mmol scale. The peptide was obtained as a white solid in an 18% yield after purification using RP-HPLC.

HRMS (ESI+) calcd for $C_{76}H_{101}N_{19}O_{17}$ $[M+2H]^{2+}$ 775.8806, found 775.8786. HPLC: 12.0 min 96.2% (20 min gradient), 23.7 min 97.6% (30 min gradient).

16



Scheme 84. Sequence of compound 16.

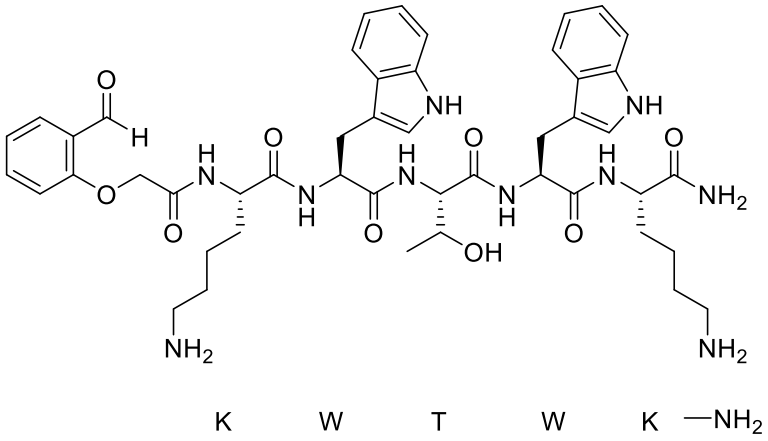
0.1 mmol of 2-Cl-Trityl chloride resin (0.74 mmol/g loading) were pre-swelled following the general procedure. The resin was then washed with DMF and treated with 3 mL of a solution of 3% hydrazine hydrate solution (~80%) in DMF (20 equiv. hydrazine) for 30 min. The resin was washed 4 x DMF and this process was repeated.

Capping was then performed with 4 mL of a 10 % MeOH in DMF solution for 30 min and the resin was then washed 4 x DMF.

In a 15 mL plastic tube, Fmoc-Glu(^{t}Bu)-OH (5 equiv.), HATU (4.5 equiv.) and DMF (3 mL) were mixed. Then DIPEA (6 equiv.) was added and the mixture was preactivated for 5 min. The solution was added to the resin and left to react on a rotary mixer for 2.5 h. The resin was then washed 4 x DMF and put on the peptide synthesiser to continue the synthesis following the general procedure for microwave-assisted SPPS. Cleavage from the resin was done following the general procedure. **16** was obtained as a white solid in a 20% yield after purification using RP-HPLC.

HRMS (ESI+) calcd for $C_{34}H_{44}N_9O_9$ $[M+H]^+$ 722.3257, found 722.3264. HPLC: 22.1 min 90.8% (30 min gradient), 20.3 min 88.8% (50 min gradient).

17



Scheme 85. Sequence of compound 17.

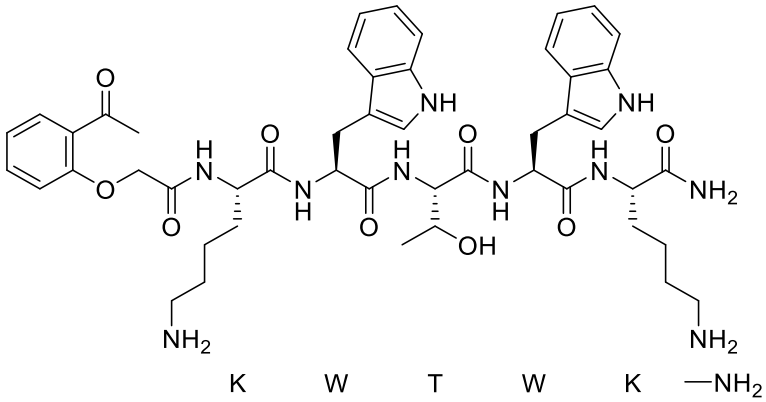
The synthesis was started following the general procedure for microwave-assisted SPPS, using TentaGel S RAM resin (0.24 mmol/g loading) in a 0.05 mmol scale. After deprotection of Lys, the resin was removed from the synthesiser and transferred to a 20 mL reaction vessel.

In a 15 mL plastic tube, 2-formylphenoxyacetic acid (5 equiv.), Oxyma Pure (4.5 equiv.) and DMF (3 mL) were mixed. DIPEA (6 equiv.) and DIC (4.5 equiv.) were then added and the mixture was mixed for 5 min to preactivate. The above solution was added onto the resin and it was mixed for 2.5 h.

The resin was then washed 4 x DMF and 4 x DCM and the cleavage was performed following the TIPS-free general procedure. After purification *via* RP-HPLC, **17** was obtained as a white solid in an 11% yield.

HRMS (ESI+) calcd for $C_{47}H_{61}N_{10}O_9$ $[M+H]^+$ 909.4618, found 909.4582. HPLC: 26 min 92.2% (30 min gradient), 23.4 min 92.4% (50 min gradient).

18



Scheme 86. Sequence of compound 18.

18 was synthesised following the general procedure for microwave-assisted SPPS, using TentaGel S RAM resin (0.24 mmol/g loading) in a 0.05 mmol scale. After deprotection of Lys, the resin was removed from the synthesiser and transferred to a 20 mL reaction vessel.

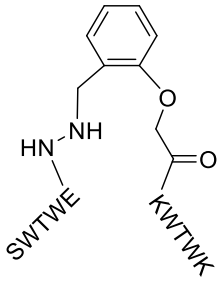
In a 15 mL plastic tube, 2-acetylphenoxyacetic acid (5 equiv.), Oxyma Pure (4.5 equiv.) and DMF (3 mL) were mixed. DIPEA (6 equiv.) and DIC (4.5 equiv.) were then added and the mixture was mixed for 5 min to preactivate. The above solution was added onto the resin and it was mixed for 2.5 h.

The resin was then washed 4 x DMF and 4 x DCM and the cleavage was performed following the general procedure. After purification *via* RP-HPLC, **18** was obtained as a white solid in a 41% yield.

HRMS (ESI+) calcd for $C_{48}H_{64}N_{10}O_9$ $[M+2H]^{2+}$ 462.2423, found 462.2443. HPLC: 13.3 min 97.7% (20 min gradient), 26.3 min 97.0% (30 min gradient).

14

The conditions for the reductive amination step were adapted from a literature procedure.[146]

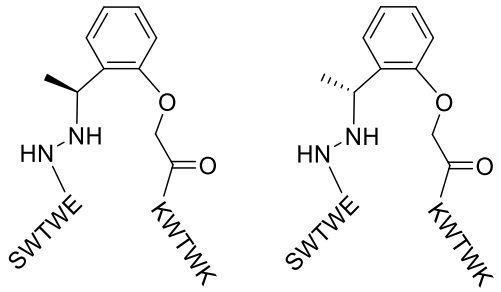


Scheme 87. Sequence of conjugate 14.

1.7 µmol of **16** and **17** were mixed as aqueous solutions, then lyophilised and re-dissolved in MeOH/AcOH 1:1 (1.5 mM) in a microwave vial. 10 equiv. $NaBH_3CN$ were then added and the reaction was stirred for 15 min. After purification *via* RP-HPLC, **14** was obtained as a white solid in a 36% yield.

HRMS (ESI+) calcd for $C_{81}H_{105}N_{19}O_{17}$ $[M+2H]^{2+}$ 807.8962, found 807.8983. HPLC: 12.7 min 96.2% (20 min gradient), 25.7 min 97.7% (30 min gradient).

(*S*)-15 and (*R*)-15



(*S*)-15 (*R*)-15

Scheme 88. Sequence of conjugates (*S*)-15 and (*R*)-15.

1.7 µmol of **16** and **18** were lyophilised and then dissolved in MeOH/AcOH 1:1 (1.5 mM) in a microwave vial. 10 equiv. $NaBH_3CN$ were then added and the reaction was stirred for 10 min. After purification *via* RP-HPLC, **(*S*)-15** and **(*R*)-15** were obtained as white solids in a 43% overall yield, with a proportion of 36% **(*S*)-15** to 64% **(*R*)-15**.

(*S*)-15

HRMS (ESI+) calcd for $C_{82}H_{107}N_{19}O_{17}$ $[M+2H]^{2+}$ 814.9041, found 814.9064. HPLC: 12.8 min 95.8% (20 min gradient), 25.9 min 96% (30 min gradient).

(*R*)-15

HRMS (ESI+) calcd for $C_{82}H_{107}N_{19}O_{17}$ $[M+2H]^{2+}$ 814.9041, found 814.9067. HPLC: 12.9 min 97.3% (20 min gradient), 26.6 min 97.2% (30 min gradient).

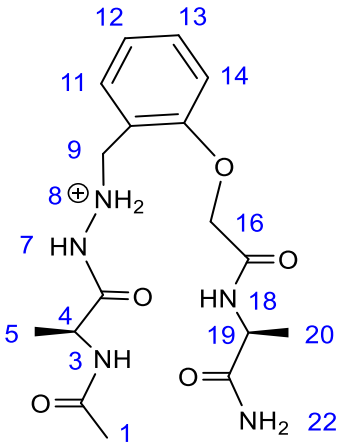
7.12.1 NMR analysis

The NMR experiments were performed in AcOH-d_4 buffer (10 mM, pH 5.5) with 5% D_2O, using DSS as an internal standard. The temperature of the experiments for **TrpZip1**, **TrpZipGG**, **14** and the tripeptide equivalents was of 288 K, and that of **(*S*)-15** and **(*R*)-15** was of 298 K. WATERGATE was used to suppress the water resonance. COSY, TOCSY, NOESY, ROESY and HSQC experiments were used to assign the resonances. Mixing times of 100, 200 or 300 msec were used for the NOESY and ROESY spectra. The chemical shifts (ppm) for all residues in each peptide are presented in this section.

7.12.1.1 Tripeptide mimic 19

Chemical shifts at 288 K:

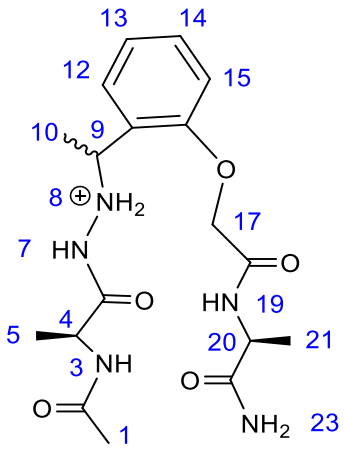
Resonance	Chemical shift (ppm)
1	1.82
3	8.08
4	4.0
5	1.06
9	4.64
11	6.81
12	7.26
13	6.92
14	7.16
16	4.04
18	8.62
19	4.27
20	1.28
22	7.55, 6.97



7.12.1.2 Tripeptide mimic 20a

Chemical shifts at 288 K:

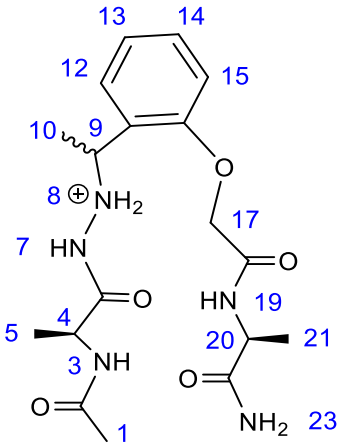
Resonance	Chemical shift (ppm)
1	1.76
3	7.96
4	3.94
5	1.0
9	4.44
10	1.29
12	7.24
13	6.94
14	7.19
15	6.8
17	4.59
19	8.52
20	4.27
21	1.3
23	7.57, 6.97



7.12.1.3 Tripeptide mimic 20b

Chemical shifts at 288 K:

Resonance	Chemical shift (ppm)
1	1.8
3	8.02
4	3.96
5	1.02
9	4.51
10	1.33
12	7.23
13	6.94
14	7.22
15	6.81
17	4.6
19	8.55
20	4.28
21	1.29
23	7.56, 6.97



7.12.1.4 TrpZip1

Chemical shifts at 288 K (ppm):

	1Ser	**2Trp**	**3Thr**	**4Trp**	**5Glu**	**6Gly**	**7Asn**
H		8.97	9.58	8.95	8.38	8.2	8.16
H^{α}	3.61	5.25	4.05	4.64	4.41	3.53, 3.82	4.03
H^{β}	3.86	3.06, 3.19	4.05	2.11, 2.97	1.80, 1.95		2.82
H^{γ}			1.17		2.09, 2.17		
H^{δ}		7.39		7			
H^{ε}		7.32, 10.27		5.54, 9.85			
H^{ζ}		6.62, 7.43		6.48, 7.22			
H^{η}		7.14		6.94			
C^{α}	57.12	57.05		56.32	54.8	46.96	52.98
C^{β}	62.32	30.87	71.49	28.47	30.47		38
C^{γ}			20.98		34.25		
C^{δ}		127.58		126.87			
C^{ε}		124.62		120.15			
C^{ε}		121.94		113.91, 121.16			
C^{η}		123.83		124.17			

	8Lys	**9Trp**	**10Thr**	**11Trp**	**12Lys**
H	6.59	8.53	9.75	9.04	7.67
H^{α}	4.22	5.19	4.06	4.33	4.2
H^{β}	1.70, 1.78	2.96, 3.30	4.06	2.09, 2.80	1.42, 1.53
H^{γ}	1.14, 1.29		1.21		1.16, 1.25
H^{δ}	1.65	7.21		6.86	1.55
H^{ε}		7.49, 9.84		5.39, 10.10	
H^{ζ}		7.12, 7.26		6.62, 7.43	
H^{η}		7.33		7.14	
C^{α}	54.99	56.85		56.79	54.09
C^{β}	34.46	30.03	71.61	28.49	33.43
C^{γ}	22.83		21.12		24.28
C^{δ}	29.08	127.61		127.61	28.62
C^{ε}		120.04		120.59	
C^{ε}		121.95, 122.43		114.62, 120.35	
C^{η}		119.83		123.83	
NH_2 cap					6.76, 7.47

7.12.1.5 TrpZipGG

Chemical shifts at 288 K (ppm):

	1Ser	**2Trp**	**3Thr**	**4Trp**	**5Glu**	**6Gly**	**7Gly**
H		8.98	9.59	8.92	8.51	8.95	7.92
H^{α}	4.92	5.27	4.93	4.73	4.42	3.74	3.23, 3.84
H^{β}		3.07	4.05	2.10, 3.00	1.86, 1.94		
H^{γ}			1.17		2.11, 2.17		
H^{δ}		7.44		6.93			
H^{ε}		7.12, 10.36		5.73, 9.95			
H^{ζ}		6.63, 7.14		6.60, 7.23			
H^{η}		7.43		6.98			
C^{α}		57.01		56.17	55.45	46.65	45.07
C^{β}		30.95	71.51	28.38	32.06		
C^{γ}			20.98		35.41		
C^{δ}		127.59		126.98			
C^{ε}		121.89		119.95			
C^{ε}		123.81		113.59, 121.35			
C^{η}		121.86		124.54			

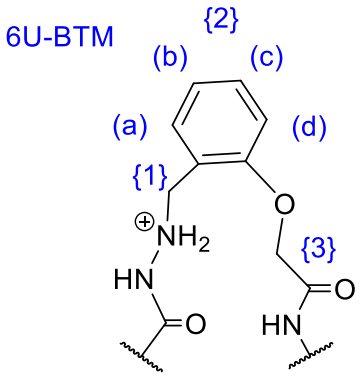
	8Lys	**9Trp**	**10Thr**	**11Trp**	**12Lys**
H	6.81	8.67	9.88	9.07	7.66
H^{α}	4.28	5.19	4.93	4.29	4.21
H^{β}	1.73	2.95, 3.29	4.08	2.03, 2.77	1.41, 1.52
H^{γ}	1.19, 1.27		1.22		1.18, 1.25
H^{δ}	1.67	7.28		6.84	1.54
H^{ε}	3.03	7.44, 9.93		5.31, 10.09	2.81, 2.86
H^{ζ}	7.58	7.25, 7.44		6.62, 7.13	
H^{η}		7.32		7.43	
C^{α}	54.31	57.07		56.7	53.79
C^{β}	35.11	29.67	71.68	28.39	33.3
C^{γ}	23.16		20.87		24.27
C^{δ}	29.03	127.68		127.61	28.44
C^{ε}	42.05	120.04		120.51	41.78
C^{ε}		114.85, 122.42		120.28	
C^{η}		124.6		114.63	
NH_2 cap					6.72, 7.45

7.12.1.6 Conjugate 14

Chemical shifts at 288 K (ppm):

	1Ser	2Trp	3Thr	4Trp	5Glu
H		8.95	9.5	8.83	8.69
H^{α}	3.57	5.2	4.8	4.73	4.04
H^{β}	3.57, 3.83	3.02	4.04	2.3, 2.82	1.62
H^{γ}			1.12		1.88, 2.03
H^{δ}		7.42		7.05	
H^{ε}		7.43, 10.33		5.90, 9.96	
H^{ζ}		7.31, 7.37		6.97, 7.03	
H^{η}		7.24		6.63	

		6U-BTM
{1}	(a)	3.73
	(b)	3.33
{2}	(a)	7.33
	(b)	7.04
	(c)	7.19
	(d)	6.82
{3}	(a)	4.56
	(b)	4.51

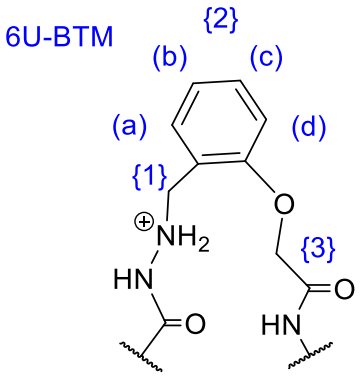


	7Lys	8Trp	9Thr	10Trp	11Lys
H	7.7	8.84	9.69	9.01	7.66
H^{α}	4.76	5.28	4.95	4.29	4.19
H^{β}	1.88	2.91, 3.3	4.01	2.02, 2.76	1.4, 1.51
H^{γ}	1.33		1.18		1.15, 1.23
H^{δ}	1.70	7.19		6.83	1.52
H^{ε}	3.00	7.25, 9.85		5.35, 10.09	2.84
H^{ζ}		7.02, 7.13		6.63, 7.42	
H^{η}		7.19		7.13	
NH_2 cap					6.73, 7.44

Chemical shifts at 278 K (ppm, used for the assignment of the carbons):

	1Ser	2Trp	3Thr	4Trp	5Glu
H		8.92	9.42	8.82	8.64
H^{α}	3.47	5.1	4.7	4.62	3.93
H^{β}	3.57, 3.71	2.90	3.94	2.18, 2.70	1.51
H^{γ}			1.02		1.76, 1.92
H^{δ}		7.33		6.95	
H^{ε}		10.3		7.09, 9.91	
H^{ζ}		7.33		6.93	
H^{η}		7.2			
C^{α}	57.04	under water	under water	under water	54.8
C^{β}	62.27	30.83	71.34	29.17	30.74
C^{γ}			20.86		35.58
C^{δ}		127.39		126.81	
C^{ε}					
C^{ε}		114.71		113.92	
C^{η}					

		6U-BTM
{1}	(a, b) CH_2	3.23, 3.64
	C CH_2	52.02
{2}	(a)	7.23
	(b)	6.93
	(c)	7.09
	(d)	6.73
	C (a)	134.47
	C (b)	125
	C (c)	132.57
	C (d)	115.87
{3}	(a, b) CH_2	4.41, 4.48
	C CH_2	71.17



	7Lys	**8Trp**	**9Thr**	**10Trp**	**11Lys**
H	7.58	8.86	9.67	9.01	7.57
H^{α}	4.66	5.18	4.83	4.17	4.07
H^{β}	1.78	2.80, 3.21	3.9	1.86, 2.63	1.29, 1.38
H^{γ}	1.22, 1.23		1.08		1.04, 1.12
H^{δ}	1.6	7.11		6.71	1.41
H^{ε}	2.89	7.14, 9.83		5.16, 10.02	2.73
H^{ζ}	7.56	7.1		6.52, 7.30	
H^{η}				7.01	
C^{α}	under water	under water	under water	56.55	53.85
C^{β}	35.17	30.28	71.47	28.25	33.33
C^{γ}	23.73		21.46		24.21
C^{δ}	29.04	127.48		127.49	28.45
C^{ε}	41.79	122.38			41.66
C^{ε}		114.76		114.42, 120.16	
C^{η}				123.71	
NH_2 cap					6.65, 7.38

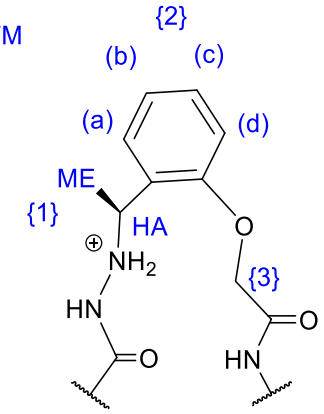
7.12.1.7 Conjugate (*S*)-15

Chemical shifts at 298 K (ppm):

	1Ser	**2Trp**	**3Thr**	**4Trp**	**5Glu**
H		8.96	9.57	8.93	8.53
H^{α}	3.6	5.27	4.9	4.71	3.95
H^{β}	3.89	3.12	4.1	2.34, 2.86	1.31, 1.37
H^{γ}			1.18		1.83, 1.97
H^{δ}		7.5		7.17	
H^{ε}		7.32, 10.37		5.90, 10.24	
H^{ζ}		7.47, 7.53		6.72, 7.45	
H^{η}		7.39		7.14	
C^{α}	57.28	57.2	under water	under water	56
C^{β}	-	30.92	at 288 K 71.46	29.52	30.73
C^{γ}			20.87		35.58
C^{δ}		127.68		127.01	
C^{ε}		122.55		-	
C^{ε}		115.04, 120.20		114.46, 121.22	
C^{η}		124.72		124.53	

		6(*S*)-BTM
{1}	ME	1.48
	C ME	20.32
	HA	4.16
	C HA	under water
{2}	(a)	7.45
	(b)	7.17
	(c)	7.09
	(d)	6.76
	C (a)	131.24
	C (b)	132.05
	C (c)	125.53
	C (d)	116.24
{3}	CH_2	4.60, 4.68
	C CH_2	at 288 K 75.68

6(*S*)-BTM



	7Lys	**8Trp**	**9Thr**	**10Trp**	**11Lys**
H	7.17	8.9	9.71	8.97	7.74
H^{α}	4.72	5.33	5.02	4.38	4.28
H^{β}	1.96	3.03, 3.40	4.07	2.15, 2.87	1.48, 1.60
H^{γ}	1.28, 1.37		1.23		1.24, 1.31
H^{δ}	1.77	7.36		6.91	1.62
H^{ε}	3.06	7.45, 10.03		5.48, 10.12	2.92
H^{ζ}		7.15, 7.33		6.70, 7.5	
H^{η}		7.25		7.21	
C^{α}	56.13	56.58	under water	56.85	54.21
C^{β}	35.11	30.77	at 288 K 71.50	28.63	33.6
C^{γ}	23.61		21.58		24.45
C^{δ}	29.3	127.79		127.76	28.74
C^{ε}	42.06	-		120.68	41.97
C^{ε}		114.97, 121.98		114.63, 120.60	
C^{η}		124.44		123.96	
NH_2 cap					6.78, 7.44

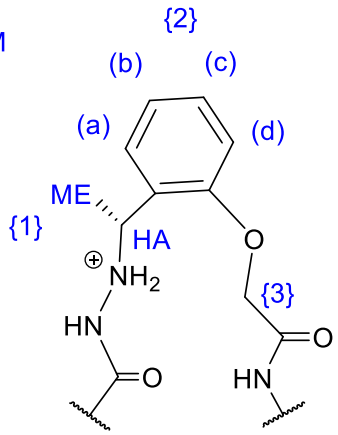
7.12.1.8 Conjugate (*R*)-15

Chemical shifts at 288 K (ppm):

	1Ser	**2Trp**	**3Thr**	**4Trp**	**5Glu**
H		8.83	9.15	8.49	8.66
H^{α}	3.6	5.06	4.62	4.85	4.09
H^{β}	3.79	2.61, 2.85	3.98	2.43, 2.81	1.68, 1.84
H^{γ}			1.03		1.99
H^{δ}		7.36		7	
H^{ε}		10.27		6.45, 9.97	
H^{ζ}		7.12, 7.38		6.84, 7.19	
H^{η}		7.25		7.02	
C^{α}	54.5	under water	57.7	under water	51.37
C^{β}	59.78	28.35	68.44	26.67	29.45
C^{γ}			18.02		33.08
C^{δ}		124.63		124.57	
C^{ε}				117.66	
C^{ε}		111.95, 123.15		111.47, 118.86	
C^{η}		121.78		121.47	

		6(*R*)-BTM
{1}	ME	1.36
	C ME	17.19
	HA	4.69
	C HA	under water
{2}	(a)	7.4
	(b)	7.12
	(c)	7.21
	(d)	6.86
	C (a)	127.89
	C (b)	117.75
	C (c)	129.21
	C (d)	114.79
{3}	CH_2	4.37, 4.49
	C CH_2	69.18

6(*R*)-BTM



	7Lys	**8Trp**	**9Thr**	**10Trp**	**11Lys**
H	8.13	8.99	9.55	8.88	7.6
H^{α}	4.74	5.02	4.78	4.15	4.11
H^{β}	1.91	2.82, 3.24	3.95	1.91, 2.68	1.32, 1.44
H^{γ}	1.36, 1.38		1.13		1.09, 1.16
H^{δ}	1.72	7.02		6.77	1.46
H^{ε}	3	6.69, 9.61		5.39, 10.03	2.75, 2.80
H^{ζ}		6.80, 7.14		6.65, 7.38	
H^{η}		7.08		7.1	
C^{α}	52.06	under water	under water	53.93	51.23
C^{β}	32.02	27.2	69.01	25.65	30.47
C^{γ}	21.5		18.22		21.51
C^{δ}	26.44	124.79		124.88	25.71
C^{ε}	39.37	117.59		117.95	39.05
C^{ε}		112.08, 118.71		111.84, 117.67	
C^{η}		121.61		121.18	
NH_2 cap					6.64, 7.27

7.12.2 Circular dichroism

CD experiments were run in potassium phosphate buffer (20 mM, pH 7) at a 0.05 mg/mL concentration for far UV (30 μM), and at a 1 mg/mL concentration for near UV (0.60 mM). A 0.1 cm cell was used.

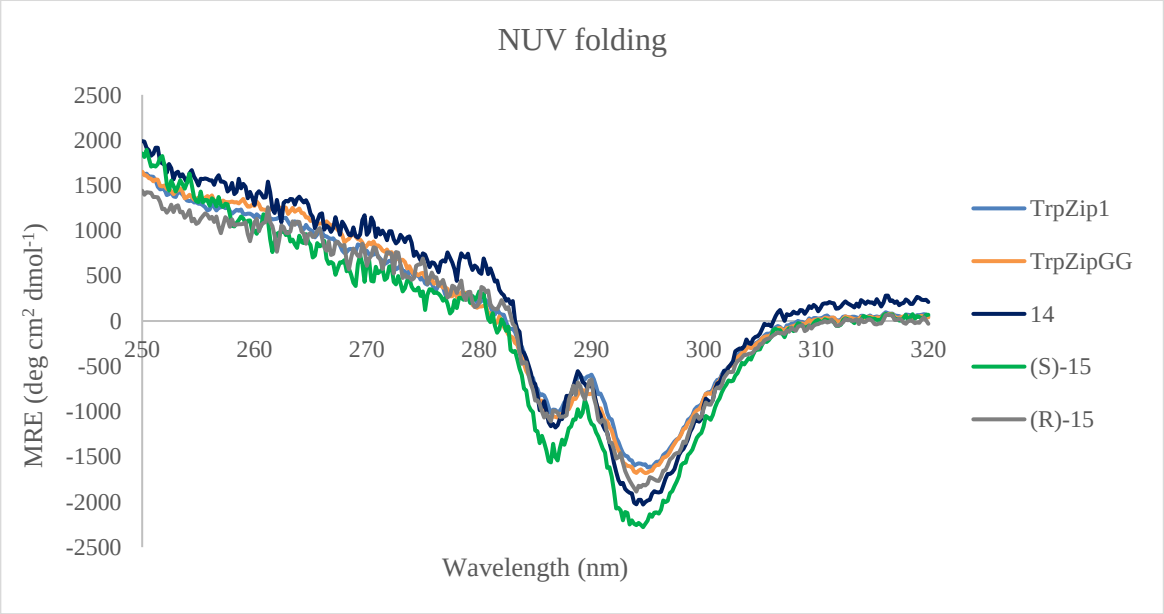
Near UV experiments were run from 320 to 250 nm, at a 20 nm/min speed with 1 sec response. Far UV experiments were run from 260 to 185 nm, at a 50 nm/min speed with a 1 sec response. Thermal denaturation experiments were performed in the far UV from 5 to 80 °C, monitoring every 5 °C. The CD was also monitored at 228 nm every 1 °C.

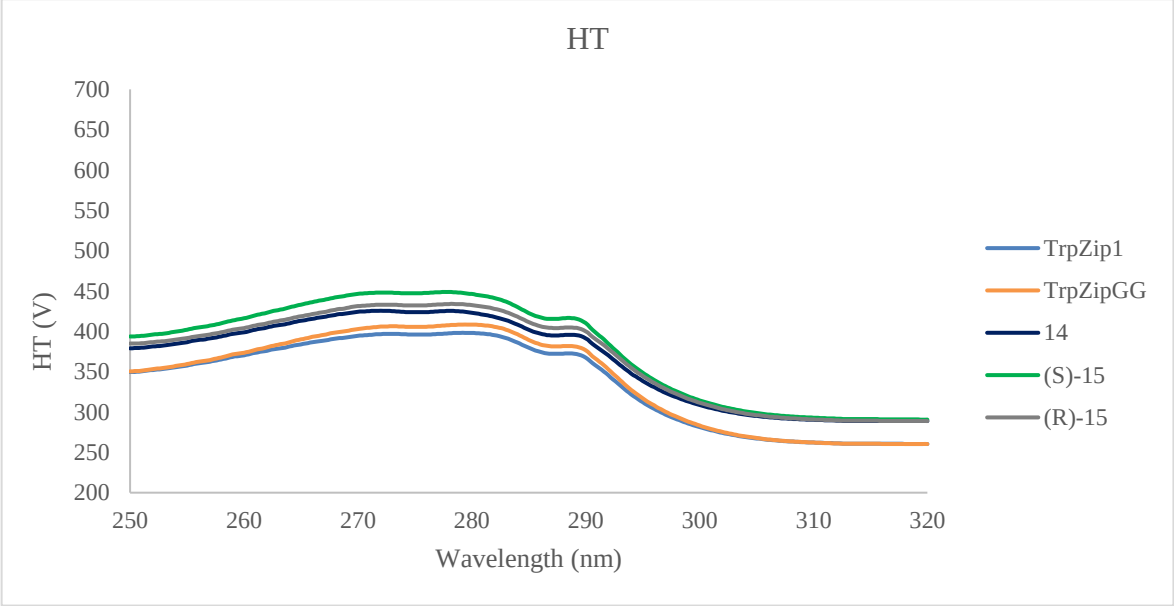
The calculation of the mean residue ellipticity (MRE, deg cm^2 $dmol^{-1}$) was performed as previously described:[365]

$$\text{MRE} = \frac{100 * \theta}{CMR * l} = \frac{100 * \theta}{C * N * l}$$

Where θ are the degrees of ellipticity in °, CMR is the mean residue concentration (concentration, C, in M * number of peptide bonds, N) and l is the pathlength in cm.

7.12.2.1 Near UV spectra





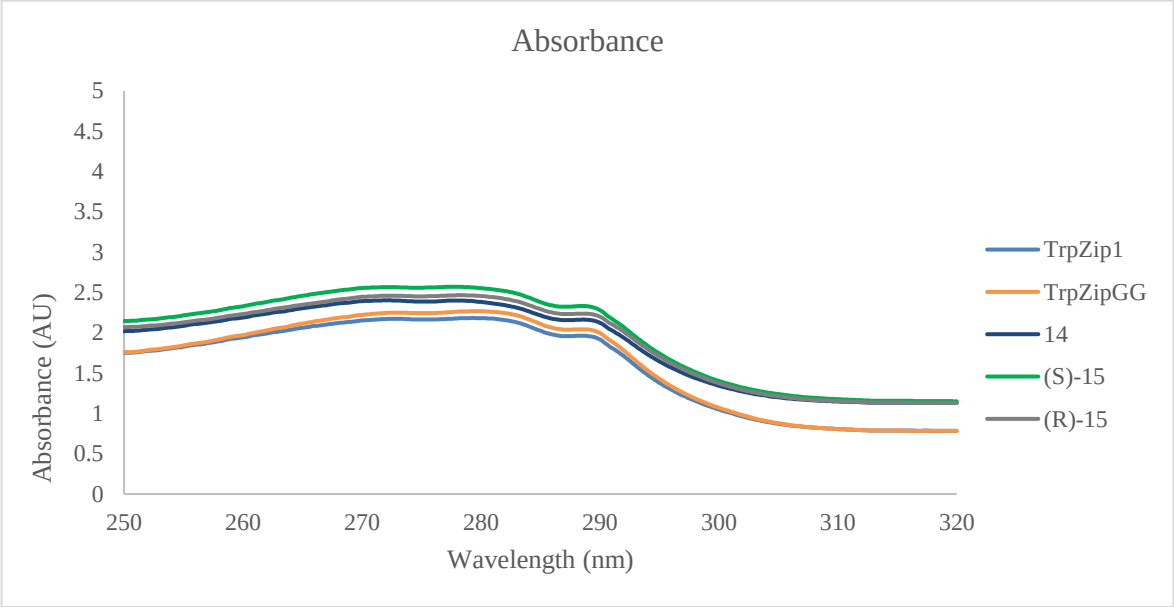
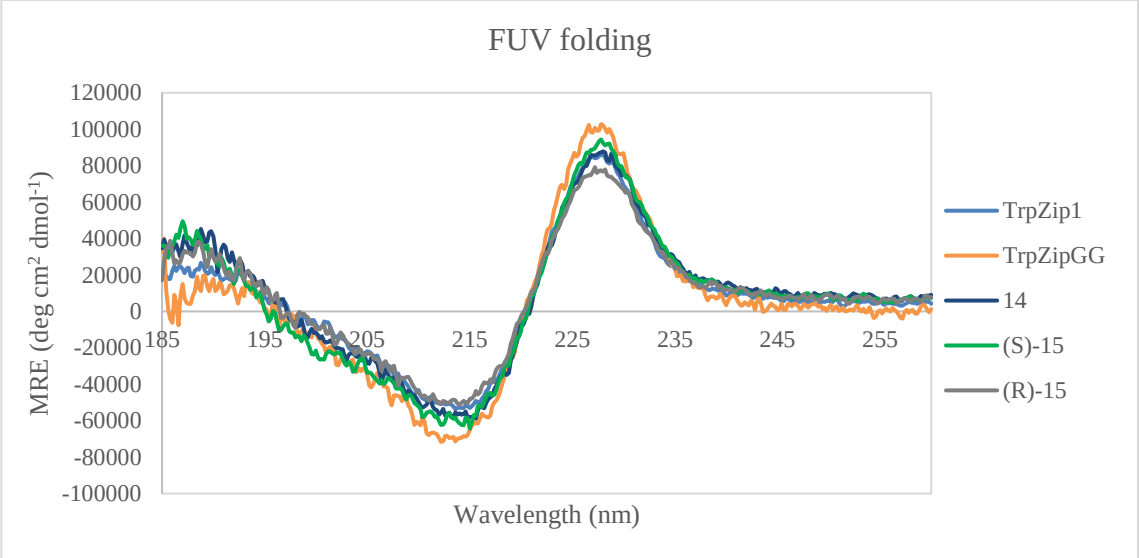
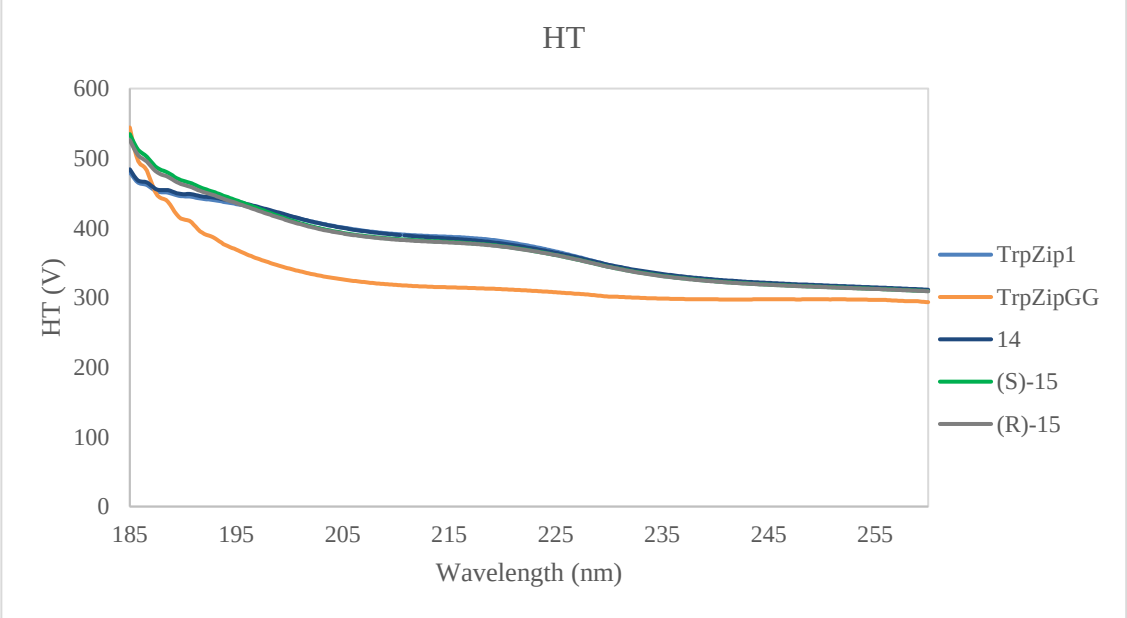


Figure 187. Near UV spectra for the peptides under study, showing the HT and absorbance graphs. Experiments were run at a 1 mg/mL concentration in potassium phosphate buffer (20 mM, pH 7), in a 0.1 cm cell.

7.12.2.2 Far UV spectra





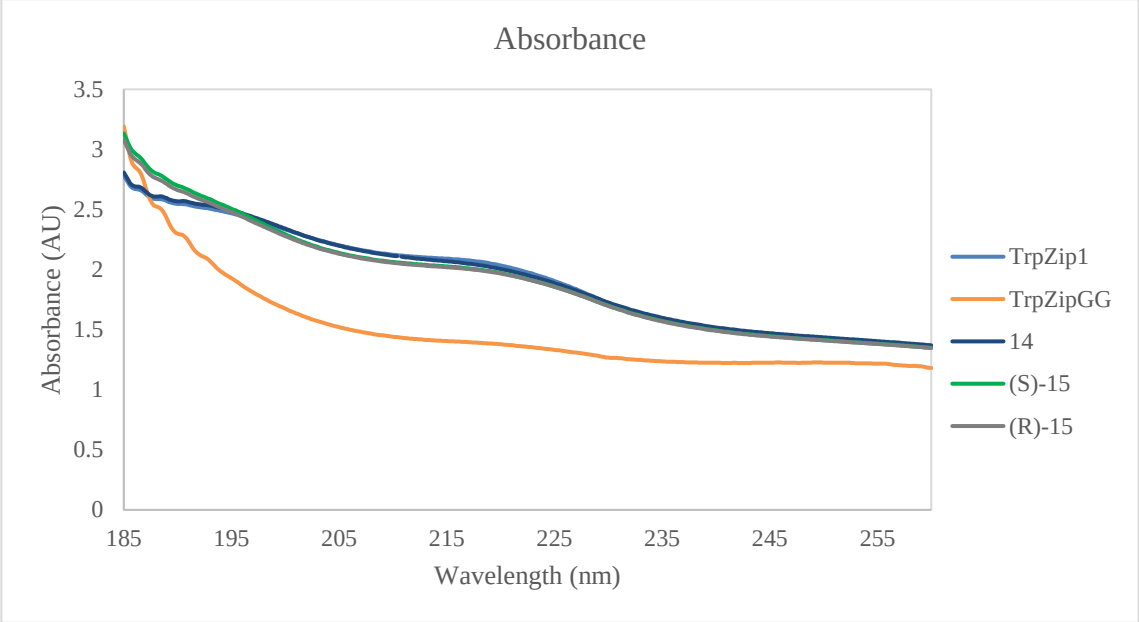


Figure 188. Far UV spectra for the peptides under study, showing the HT and absorbance graphs. Experiments were run at a 0.05 mg/mL concentration in potassium phosphate buffer (20 mM, pH 7), in a 0.1 cm cell.

7.12.2.3 Thermal denaturation

Monitoring in the far UV every 5 degrees, from 5 to 80 °C.

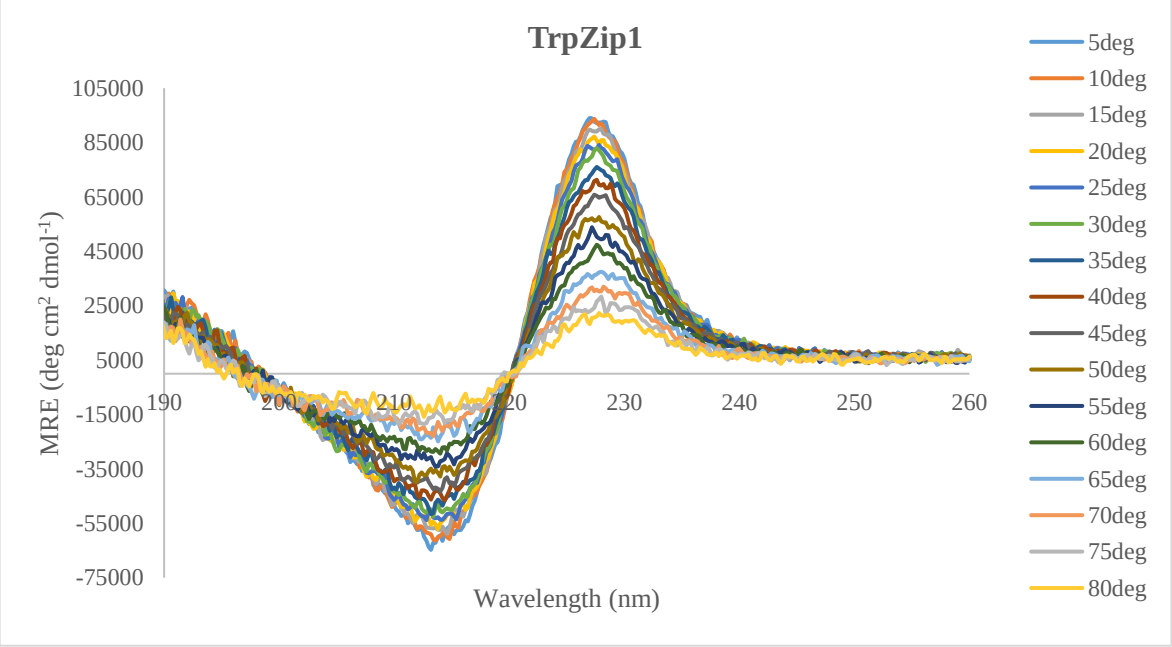


Figure 189. Thermal denaturation for TrpZip1, run at a 0.05 mg/mL concentration in potassium phosphate buffer (20 mM, pH 7), in a 0.1 cm cell.

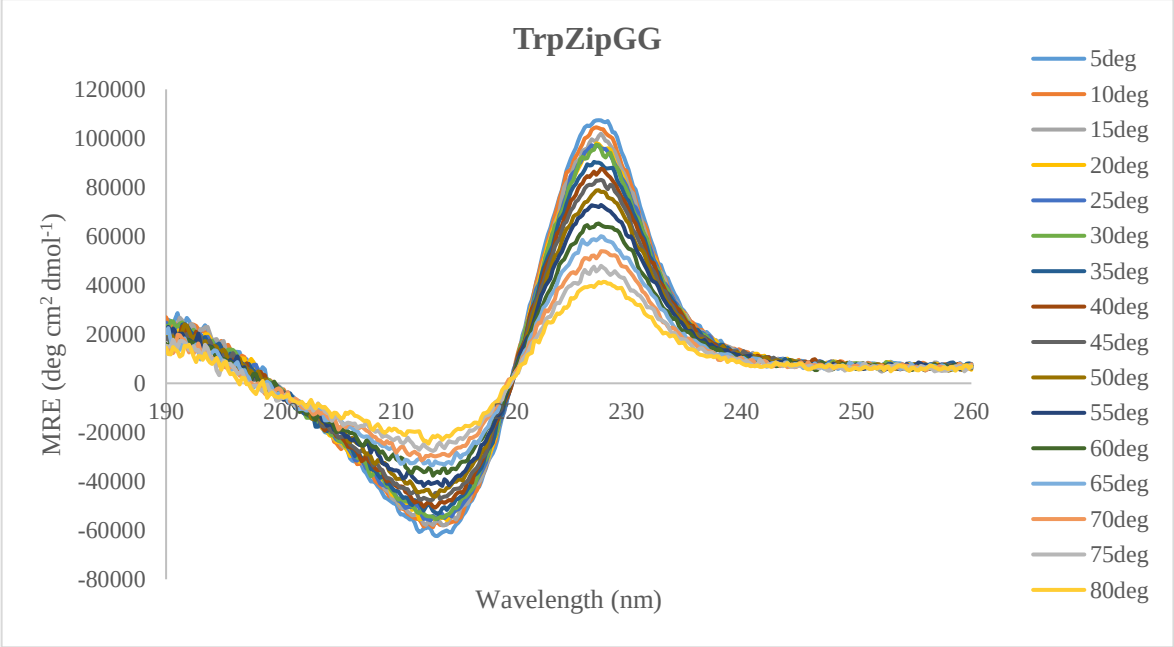


Figure 190. Thermal denaturation for TrpZipGG, run at a 0.05 mg/mL concentration in potassium phosphate buffer (20 mM, pH 7), in a 0.1 cm cell.

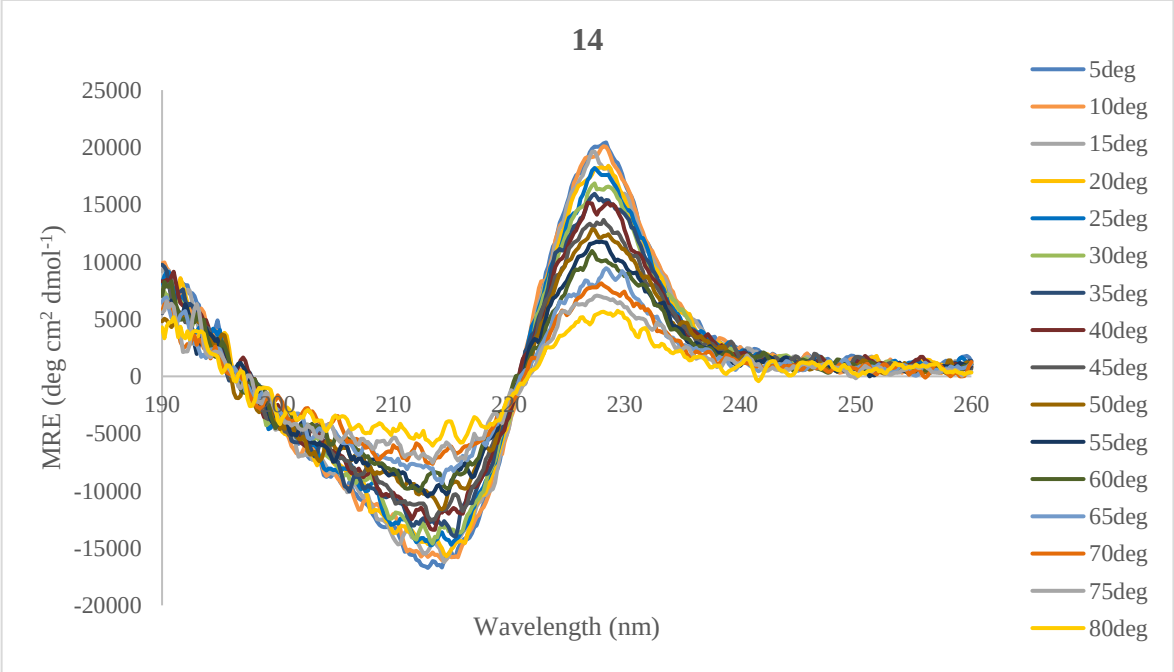


Figure 191. Thermal denaturation for 14, run at a 0.05 mg/mL concentration in potassium phosphate buffer (20 mM, pH 7), in a 0.1 cm cell.

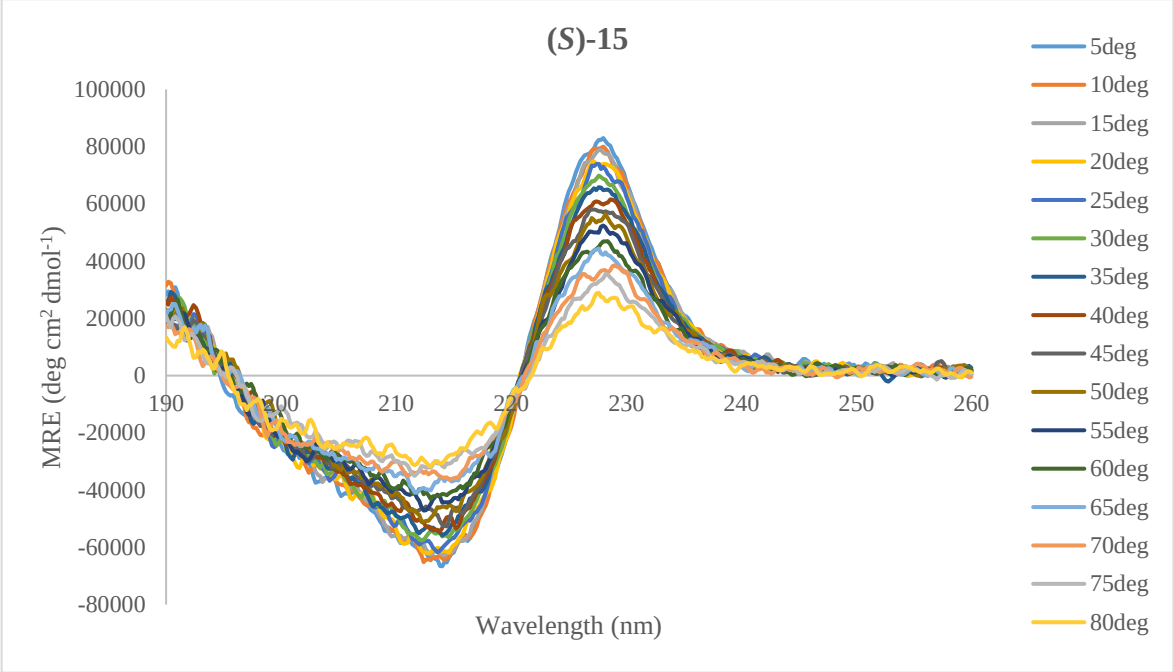


Figure 192. Thermal denaturation for (*S*)-15, run at a 0.05 mg/mL concentration in potassium phosphate buffer (20 mM, pH 7), in a 0.1 cm cell.

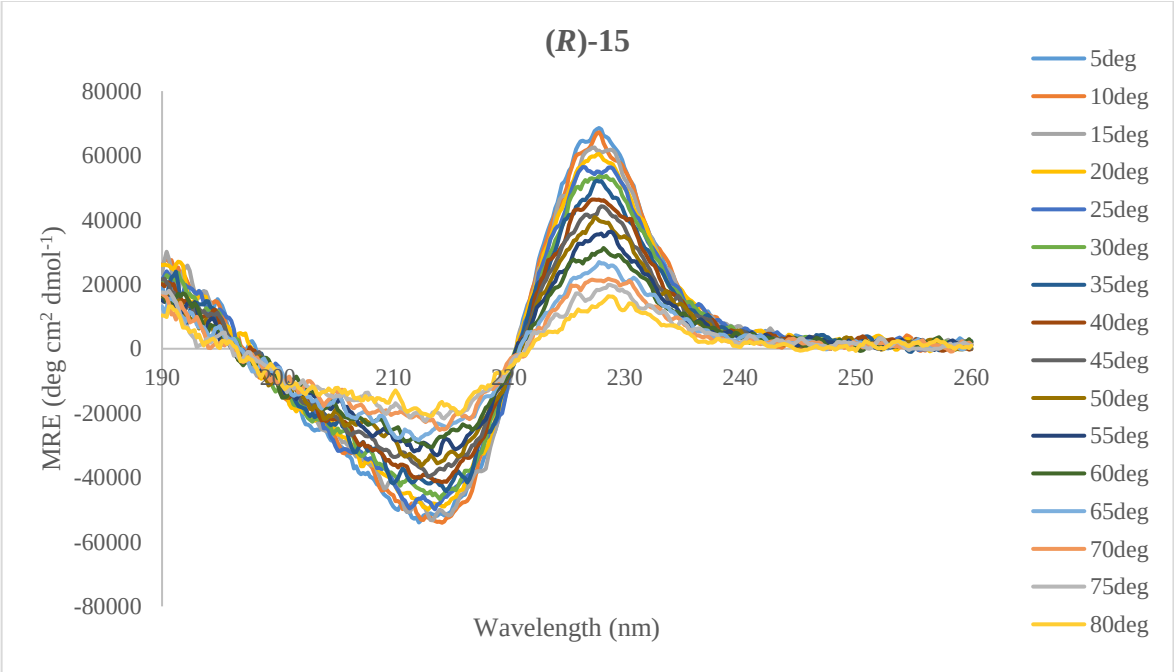


Figure 193. Thermal denaturation for (*R*)-15, run at a 0.05 mg/mL concentration in potassium phosphate buffer (20 mM, pH 7), in a 0.1 cm cell.

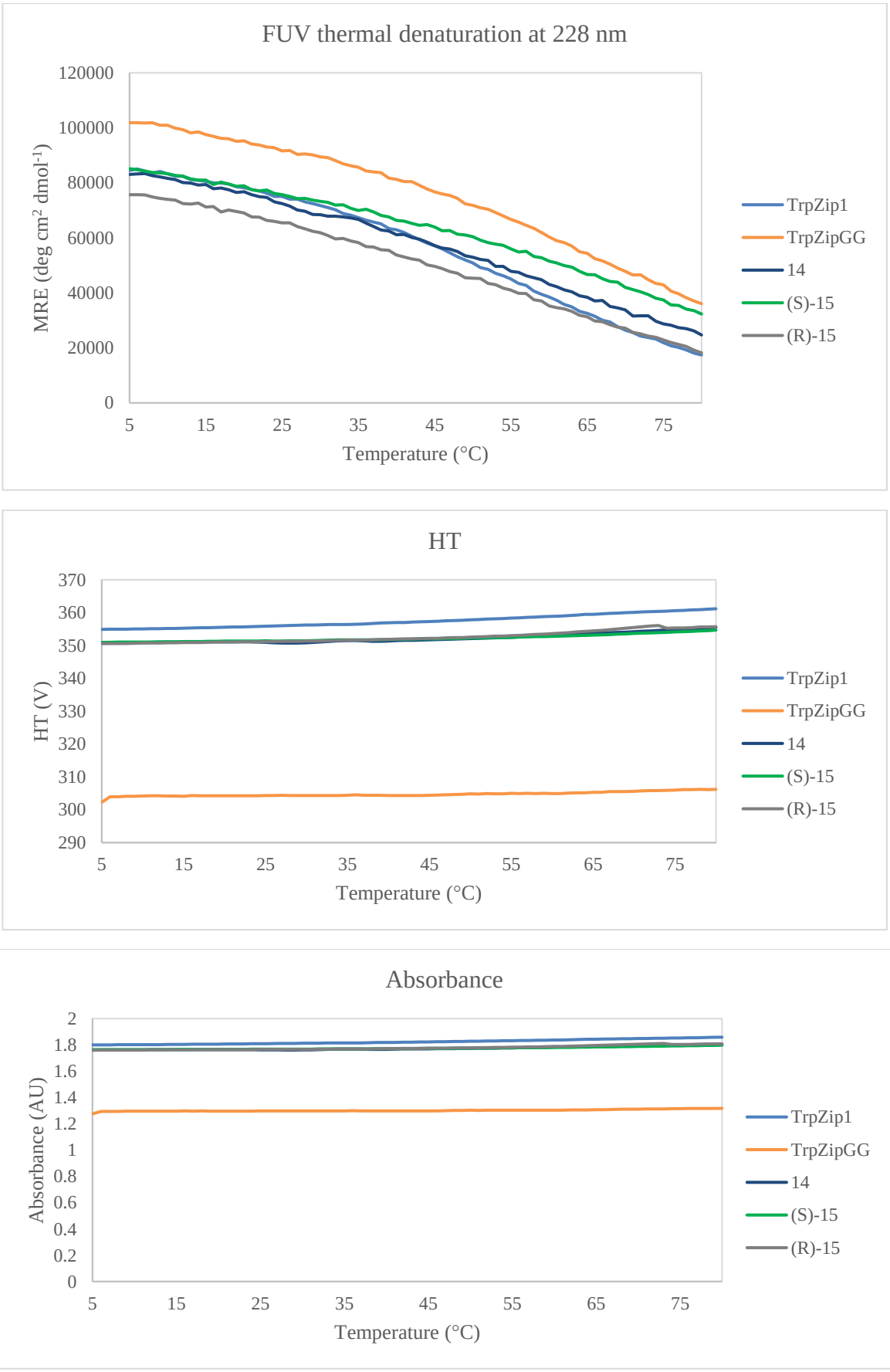
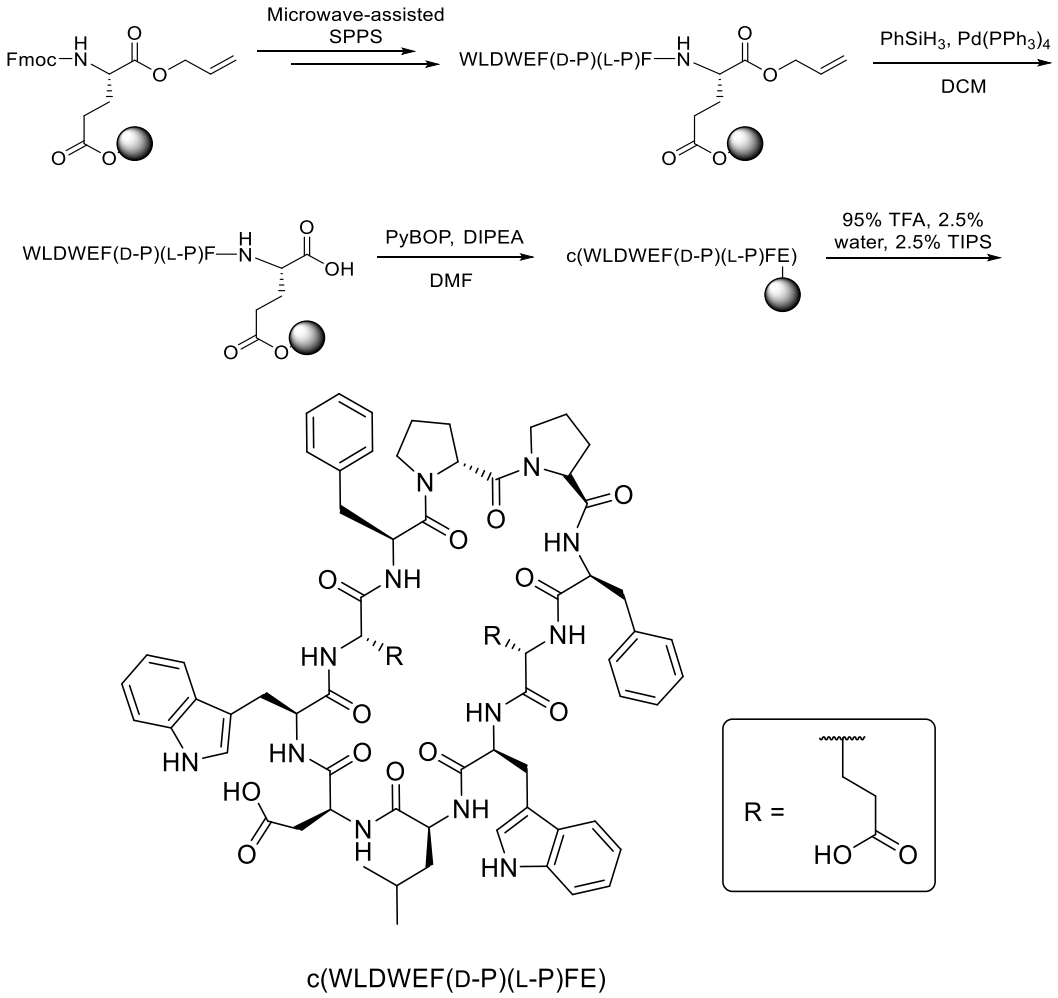


Figure 194. Thermal denaturation curves measured at 228 nm for all peptides under study, showing the HT and absorbance curves. Experiments were run at 0.05 mg/mL concentrations in potassium phosphate buffer (20 mM, pH 7), in a 0.1 cm cell.

7.13 Chapter 4 peptides

29



Scheme 89. Synthetic route and sequence of compound 29.

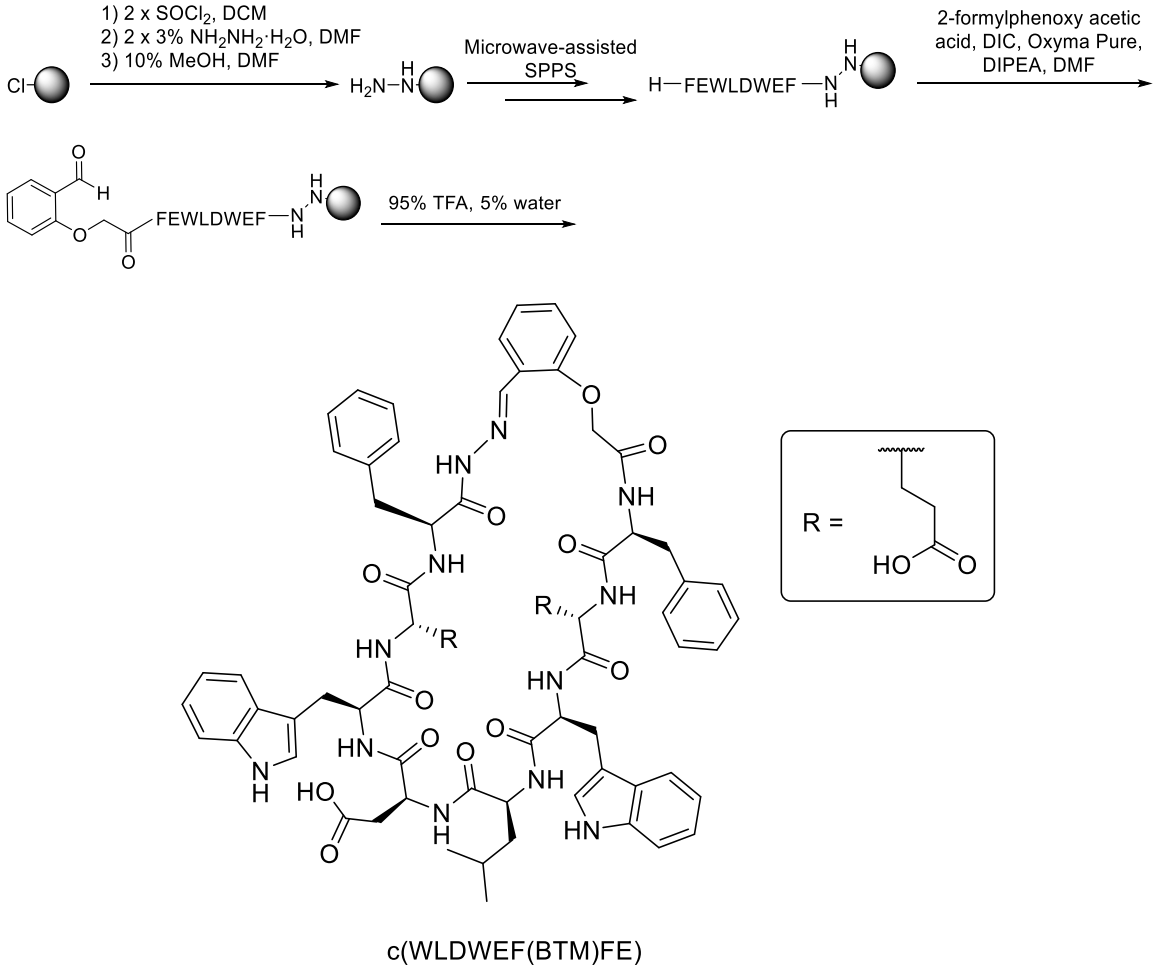
The synthesis of **29** was performed on Fmoc-Glu(Wang)-OAll resin (0.44 mmol/g, 0.1 mmol scale), using standard microwave-assisted SPPS. After deprotection of the *C*-terminal Trp residue the allyl group was deprotected: to this end, a solution of $PhSiH_3$ (24 equiv.) in 2 mL DCM was added to the resin and stirred for 2 min. Then a solution of $Pd(PPh_3)_4$ (0.25 equiv.) in 6 mL DCM was added onto the resin and mixed for 30 min. After that time the resin was washed 3 x DCM, 3 x DMF, 3 x DCM and the process was repeated.

A solution of PyBOP (5 equiv.) in DMF (6 mL) with DIPEA (6 equiv) was preactivated for 5 min and then added on the resin to perform the on-resin cyclisation. The reaction was mixed for 2.5 h at rt and the resin was then washed 3 x DMF. Cleavage from the resin was

performed following the general procedure. Purification *via* RP-HPLC afforded **29** in a 7% yield.

HRMS (ESI+) calcd for $C_{70}H_{82}N_{12}NaO_{16}$ $[M+Na]^+$ 1369.5864, found 1369.5851. HPLC: 19.2 min 97.9% (20 min gradient), 37.8 min 97.9% (50 min gradient).

30



Scheme 90. Synthetic route and sequence of compound 30.

0.1 mmol of 2-Cl-trityl chloride resin (0.8 mmol/g loading) were swelled following the general procedure. A solution of 3 % $SOCl_2$ (15 equiv.) in DCM (3 mL) was added to the resin, which was shaken for 30 min. The resin was washed 3 x DCM and this step was repeated.

The resin was then treated with a 3 % hydrazine hydrate solution (~80%, 20 equiv.) in DMF (3 mL) for 30 min, after which time the resin was washed 3 x DMF. This process was

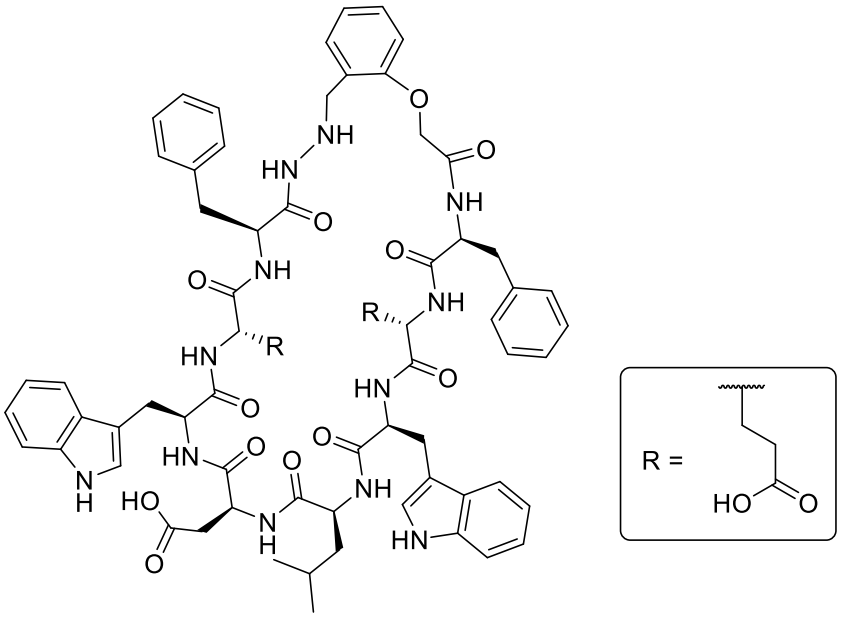
repeated twice. The resin was then capped using a 10 % MeOH/DMF solution (4 mL) for 30 min.

The peptide chain was elongated following the general procedure for microwave-assisted SPPS. After deprotection of the *C*-terminal Phe, a solution of 2-formylphenoxyacetic acid (5 equiv.), Oxyma Pure (4.5 equiv.), DIC (4.5 equiv.) and DIPEA (6 equiv.) in DMF (3 mL) (preactivated for 5 min) was added to the resin. The coupling was left for reaction for 2.5 h at rt.

Cleavage from the resin was then performed using the TIPS-free general procedure. After purification *via* RP-HPLC, **30** was obtained in a 3% yield.

HRMS (ESI+) calcd for $C_{69}H_{76}N_{12}NaO_{16}$ $[M+Na]^+$ 1351.5394, found 1351.5378. HPLC: 22.2 min 95.3% (20 min gradient), 44.9 min 90.3% (50 min gradient).

28



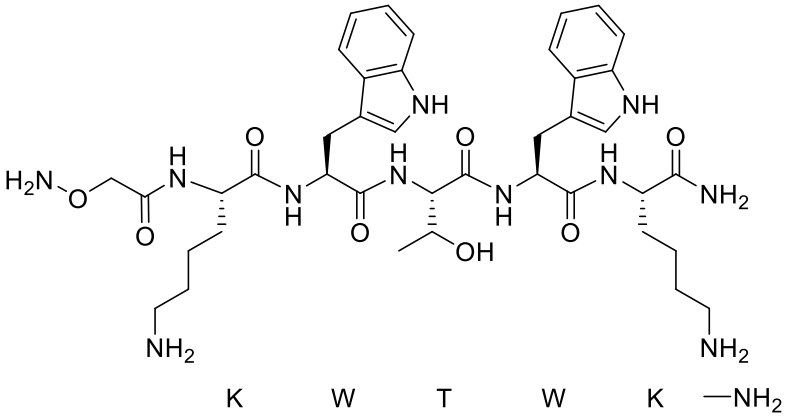
Scheme 91. Sequence of compound 28.

2 μmol of **30** were dissolved in MeOH/AcOH 1:1 (1.5 mM) in a microwave vial. 10 equiv. $NaBH_3CN$ were then added and the reaction was stirred for 15 min. **28** was obtained in an 18% yield after purification *via* RP-HPLC.

HRMS (ESI+) calcd for $C_{69}H_{78}N_{12}NaO_{16}$ $[M+Na]^+$ 1353.5551, found 1353.5525. HPLC: 17.8 min 97.3% (20 min gradient), 34.8 min 95.4% (50 min gradient).

7.14 Chapter 5 experimental

33



Scheme 92. Sequence of compound 33.

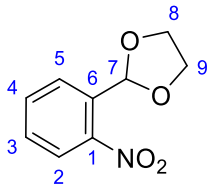
The synthesis was started following the general procedure for microwave-assisted SPPS, using Fmoc-rink amide AM resin (0.51 mmol/g loading) in a 0.25 mmol scale. After deprotection of Lys, the resin was removed from the synthesiser and transferred to a 20 mL reaction vessel.

N-hydroxyphthalimide (9.5 equiv.), bromoacetic acid (5 equiv.), DIC (4.5 equiv.) and DIPEA (6 equiv.) were dissolved in DMF (5 mL) and preactivated for 5 min. The mixture was added onto the resin and the reaction was mixed for 3 h.

After the resin was washed with 3 x DMF, DMF was added (5 mL) followed by hydrazine hydrate (80%, 2 equiv.), and the resin was mixed for 2 h. The resin was then washed 3 x DMF and 3 x DCM and the cleavage was performed following the TIPS-free general procedure. This afforded **33** as a white solid in a 7% yield after purification *via* RP-HPLC.

HRMS (ESI+) calcd for $C_{40}H_{59}N_{11}O_8$ $[M+2H]^{2+}$ 410.7269, found 410.7273. HPLC: 11.13 min 97.2% (20 min gradient), 20.6 min 96.5% (30 min gradient).

2-(1,3-Dioxolan-2-yl)benzenamine, 35

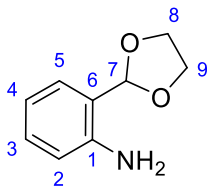


In a 100 mL flask, 2-nitrobenzaldehyde (1 equiv., 6.6 mmol) and *p*-toluenesulfonic acid monohydrate (0.2 equiv.) were dissolved in dry toluene (60 mL) under a nitrogen atmosphere. Then, dry ethylene glycol (10 equiv.) was added and the reaction mixture was stirred under reflux overnight. After cooling the crude to room temperature, it was dried *in vacuo* and then diluted with DCM and sat. $NaHCO_3$. The aqueous layer was then washed 3 x DCM and the combined organic layers were dried over $MgSO_4$.[337] After drying *in vacuo*, **35** was obtained in quantitative yield as an orange oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.90 (dd, *J* = 8.0, 1.4 Hz, 1H, Ar-H), 7.80 (dd, *J* = 8.0, 1.4 Hz, 1H, Ar-H), 7.62 (td, *J* = 8.0, 8.0, 1.4 Hz, 1H, Ar-H), 7.50 (ddd, *J* = 8.0, 8.0, 1.4 Hz, 1H, Ar-H), 6.48 (s, 1H, 7-H), 4.09 – 3.99 (m, 4H, 8-H, 9-H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 148.88 (Ar-C), 133.21 (Ar-C), 132.87 (Ar-CH), 129.66 (Ar-CH), 127.65 (Ar-CH), 124.43 (Ar-CH), 99.59 (CH), 65.34 (CH_2). HRMS (ESI+) calcd for $C_9H_9NNaO_4$ $[M+Na]^+$ 218.0424, found 218.0424.

The spectroscopic data was in good agreement with the literature.[337]

N-(2-(1,3-Dioxolan-2-yl)phenylamine, 34

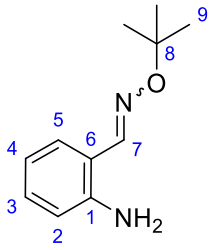


A solution of **35** (1 equiv., 1.5 mmol) in ethanol was added to a stirred suspension of Na_2S (2.5 equiv.) in ethanol (5 mL) under a nitrogen atmosphere. The mixture was stirred for 1 h and then triethylamine (0.7 equiv.) was added. After stirring for 10 min, the reaction mixture was concentrated under vacuum. The residue was diluted in Et_2O, Et_3N and H_2O, the aqueous phase was washed with Et_2O and the combined organic layers were washed with H_2O.[341] After drying the organic phases over $MgSO_4$ and concentrating them over vacuum, **34** was obtained in a 65% yield (0.16 g) as an orange oil.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.30 (dd, *J* = 7.6, 1.6 Hz, 1H, Ar-H), 7.15 (td, *J* = 7.6, 1.6 Hz, 1H, Ar-H), 6.75 (td, *J* = 7.5, 1.2 Hz, 1H, Ar-H), 6.68 (dd, *J* = 8.0, 1.2 Hz, 1H, Ar-H), 5.82 (s, 1H, 7-H), 4.20 (s, 2H, NH_2), 4.13 – 4.03 (m, 4H, 8-H, 9-H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 145.26 (Ar-C), 130.01 (Ar-CH), 127.52 (Ar-CH), 121.20 (Ar-C), 117.94 (Ar-CH), 116.47 (Ar-CH), 103.36 (CH), 64.88 (2 x CH_2). HRMS (ESI+) calcd for $C_9H_{12}NO_2$ $[M+H]^+$ 166.0863, found 166.0863.

The spectroscopic data was in good agreement with the literature.[366]

O-^{t}Bu-2-aminobenzaldehyde oxime, 36

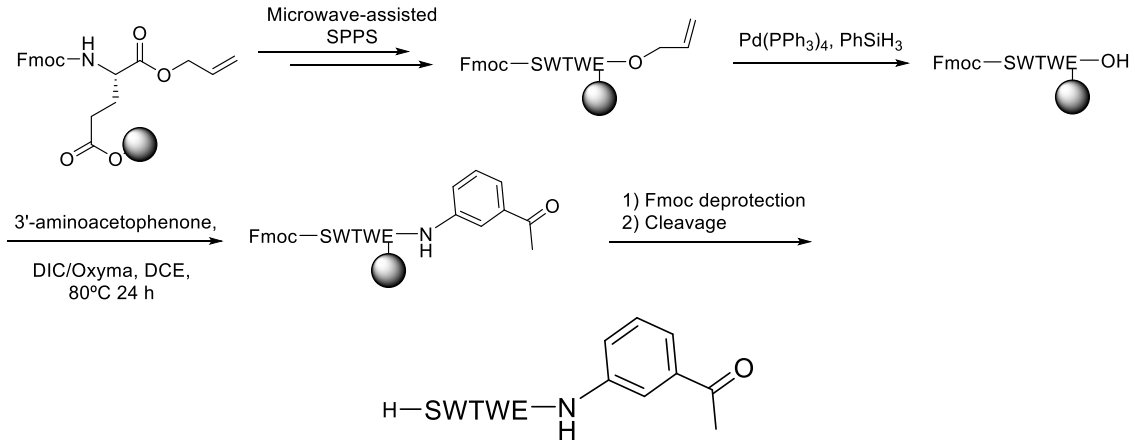


A mixture of O-^{t}Bu hydroxylamine·HCl (2 equiv.) and 2-aminobenzaldehyde (1 equiv., 0.5 mmol) in ethanol (2.5 mL) was stirred for 10 min and then NaOAc (3 equiv.) was added. The suspension was stirred at 50 °C for 6 h. Once cooled to rt, the crude was concentrated under vacuum and re-dissolved in DCM. H_2O and Et_3N were then added, and the aqueous layer was washed 3 x DCM. The combined organic phases were washed 1 x sat. NaCl solution and dried over $MgSO_4$.[343] After drying *in vacuo*, **36** was obtained as a yellow oil (82.6 mg, 43% yield).

^{1}H NMR (400 MHz, $CDCl_3$) δ 8.17 (s, 1H, 7-H), 7.13 – 7.07 (m, 2H, Ar-H), 6.70 – 6.66 (m, 2H, Ar-H), 5.56 (br s, 2H, NH_2), 1.36 (s, 9H, 9-H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 151.05 (CH), 146.19 (Ar-C), 132.09 (Ar-CH), 129.88 (Ar-CH), 116.77 (Ar-CH), 115.51 (Ar-CH), 115.37 (Ar-C), 27.58 (3 x CH_3). HRMS (ESI+) calcd for $C_{11}H_{16}N_2NaO$ $[M+Na]^+$ 215.1155, found 215.1155.

The spectroscopic data was in good agreement with the literature.[367]

38



Scheme 93. Synthetic route and sequence of compound 38.

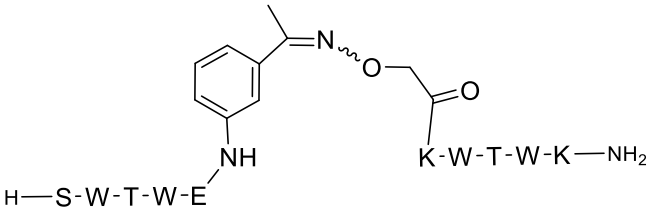
The synthesis was started following the general procedure for microwave-assisted SPPS, using 0.1 mmol of Fmoc-Glu(Wang)-OAll resin (0.44 mmol/g loading). The Ser was left Fmoc-protected and the resin was transferred to a 20 mL reaction vessel.

A solution of $PhSiH_3$ (24 equiv.) in 2 mL DCM was added onto the resin and stirred for 2 min. Then a solution of $Pd(PPh_3)_4$ (0.25 equiv.) in 6 mL DCM was added and the mixture was shaken for 30 min. The resin was washed 3 x DCM, 3 x DMF and 3 x DCM and the process was repeated.

The resin was suspended in DCE (5 mL) in a sealed microwave vial and DIC (10 equiv.) and Oxyma Pure (10 equiv.) were added. After preactivating the mixture for 5 min, 3'-aminoacetophenone (8 equiv.) was added and the reaction was stirred under reflux for 24 h. After this time, the resin was washed 3 x DMF, the Fmoc group was removed and the cleavage was performed following the TIPS-free general procedure. **38** was obtained as a white solid in a 6% yield after purification *via* RP-HPLC.

HRMS (ESI+) calcd for $C_{42}H_{49}N_8O_{10}$ $[M+H]^+$ 825.3566, found 825.3570. HPLC: 14.2 min 94% (20 min gradient), 29 min 95% (30 min gradient).

37a and 37b



Scheme 94. Sequence of compounds 37a and 37b.

38 (2.3 μmol, 1 equiv.) and **33** (1.1 equiv.) were lyophilised and then dissolved in MeOH/AcOH 1:1 (1 mM) in a microwave vial. The reaction was stirred at room temperature for 3 days. After purification *via* RP-HPLC, **37a** and **37b** were obtained in a 52% overall yield, with a proportion of 21% **37a** to 79% **37b**.

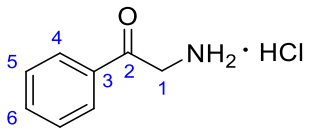
37a (*Z*-oxime)

HRMS (ESI+) calcd for $C_{82}H_{105}N_{19}O_{17}$ $[M+2H]^{2+}$ 813.8962, found 813.8960. HPLC: 13.8 min 98.1% (20 min gradient), 29 min 97.9% (30 min gradient).

37b (*E*-oxime)

HRMS (ESI+) calcd for $C_{82}H_{105}N_{19}O_{17}$ $[M+2H]^{2+}$ 813.8962, found 813.8969. HPLC: 14.3 min 99.3% (20 min gradient), 30.2 min 98.6% (30 min gradient).

2-Oxo-2-phenylethan-1-aminium chloride, 41

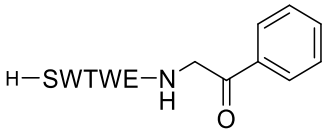


Under a nitrogen atmosphere, hexamine (1.1 equiv.) was added to a solution of 2-bromo-acetophenone (1 equiv., 5 mmol) in dichloroethane (25 mL). The reaction mixture was stirred at 60 °C for 4 h, after which time it was cooled to rt. The crude was filtered and the precipitate was suspended in EtOH (10 mL). 5 mL of conc. HCl was then added dropwise, and the mixture was stirred at rt for 18 h. The crude was then filtered and the filtrate was collected and concentrated under vacuum[346] to afford **41** as a yellow solid (1.3 g, quantitative yield).

^{1}H NMR (400 MHz, CD_3OD) δ 8.07 – 8.02 (m, 2H, Ar-H), 7.74 – 7.69 (m, 1H, Ar-H), 7.61 – 7.55 (m, 2H, Ar-H), 4.61 (s, 2H, 1-H). ^{13}C NMR (101 MHz, CD_3OD) δ 191.83 (C), 134.43 (Ar-CH), 133.65 (Ar-C), 128.80 (2 x Ar-CH), 127.85 (2 x Ar-CH), 44.78 (CH_2). HRMS (ESI+) calcd for C_8H_9NNaO $[M+Na]^+$ 158.0576, found 158.0577.

The spectroscopic data was in good agreement with the literature.[368]

40



Scheme 95. Sequence of compound 40.

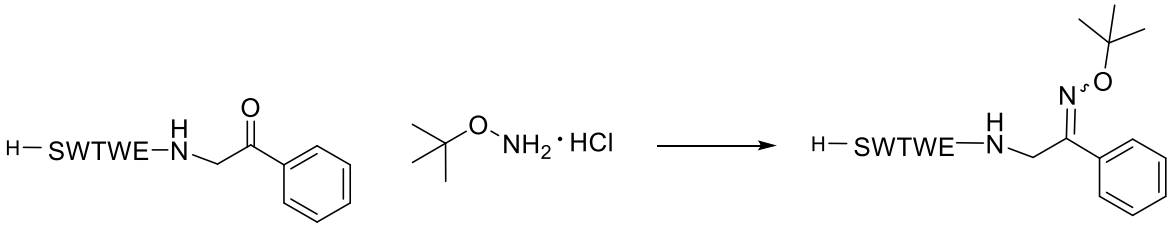
The synthesis was started following the general procedure for microwave-assisted SPPS, using 0.1 mmol of Fmoc-Glu(Wang)-OAll resin (0.44 mmol/g loading). The Ser was left Fmoc protected and the resin was transferred to a 20 mL reaction vessel.

A solution of $PhSiH_3$ (24 equiv.) in 2 mL DCM was added onto the resin and stirred for 2 min. Then a solution of $Pd(PPh_3)_4$ (0.25 equiv.) in 6 mL DCM was added and the mixture was shaken for 30 min. The resin was washed 3 x DCM, 3 x DMF and 3 x DCM and the process was repeated.

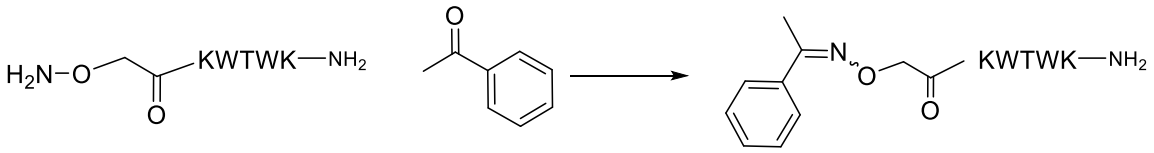
The resin was then suspended in DMF (5 mL), and DIC (1.2 equiv.) and Oxyma Pure (1.2 equiv.) were added. After preactivating for 5 min, **41** (1 equiv.) was added and the reaction was mixed at rt for 2.5 h. The resin was then washed 3 x DMF and the process was repeated. The Fmoc group was removed and the cleavage was performed following the TIPS-free general procedure. After purification using RP-HPLC, **40** was obtained as a white solid in a 36% yield.

HRMS (ESI+) calcd for $C_{42}H_{49}N_8O_{10}$ $[M+H]^+$ 825.3566, found 825.3567. HPLC: 13.9 min 98.1% (20 min gradient), 25.6 min 98.4% (50 min gradient).

Ligation tests with small molecules

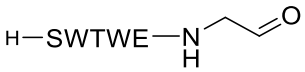


40 (1 μmol, 1 equiv.) was lyophilised and then dissolved in a NH_4OAc buffer (0.1 M, pH 4.5) and 5% DMF to a 1 mM concentration in a microwave vial. O-^{t}Bu-hydroxylamine HCl (8 equiv.) was added and the reaction mixture was stirred at rt for 5 days.



33 (1 μmol, 1 equiv.) was lyophilised and then dissolved in a NH_4OAc buffer (0.1 M, pH 4.5) and 5% DMF to a 1 mM concentration in a microwave vial. Acetophenone (8.5 equiv.) was added and the reaction mixture was stirred at rt for 1 day.

43



Scheme 96. Sequence of compound 43.

The synthesis was started following the general procedure for microwave-assisted SPPS, using 0.025 mmol of Fmoc-Glu(Wang)-OAll resin (0.44 mmol/g loading). The Ser was left Fmoc protected and the resin was transferred to a 20 mL reaction vessel.

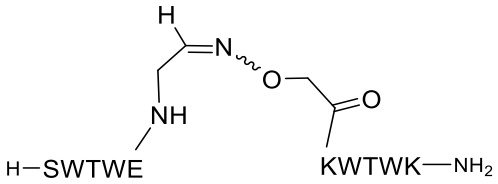
A solution of $PhSiH_3$ (24 equiv.) in 2 mL DCM was added onto the resin and stirred for 2 min. Then a solution of $Pd(PPh_3)_4$ (0.25 equiv.) in 6 mL DCM was added and the mixture was shaken for 30 min. The resin was washed 3 x DCM, 3 x DMF and 3 x DCM and the process was repeated.

The resin was then suspended in DCM (6 mL), and DIC (10 equiv.) and Oxyma Pure (10 equiv.) were added. After preactivating for 5 min, aminoacetaldehyde dimethyl acetal (8 equiv.) was added and the reaction was mixed at rt overnight. The Fmoc group was then

removed and the cleavage was performed using 85% TFA and 15% H_2O. Purification *via* RP-HPLC afforded **43** as a white solid in a 2.5% yield.

HRMS (ESI+) calcd for $C_{36}H_{45}N_8O_{10}$ $[M+H]^+$ 749.3253, found 749.3246. HPLC: 14 min 84% (20 min gradient), 28.5 min 82.5% (30 min gradient).

42a and 42b



Scheme 97. Sequence of compounds 42a and 42b.

43 (0.9 μmol, 1 equiv.) and **33** (1.2 equiv.) were lyophilised and then dissolved in MeOH/AcOH 1:1 (1 mM) in a microwave vial. The reaction was stirred at room temperature for 24 h. **42a** and **42b** were obtained in a 54% overall yield after purification *via* RP-HPLC, in a proportion of 57% **42a** to 43% **42b**.

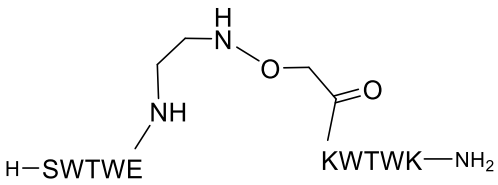
42a (*Z*-oxime)

HRMS (ESI+) calcd for $C_{76}H_{101}N_{19}O_{17}$ $[M+2H]^{2+}$ 775.8806, found 775.8797. HPLC: 12.5 min 84.2% (20 min gradient), 25.1 min 81.2% (30 min gradient).

42b (*E*-oxime)

HRMS (ESI+) calcd for $C_{76}H_{101}N_{19}O_{17}$ $[M+2H]^{2+}$ 775.8806, found 775.8804. HPLC: 12.8 min 72% (20 min gradient), 26 min 65.4% (30 min gradient).

45



Scheme 98. Sequence of compound 45.

42a (0.1 µmol, 1 equiv.) was dissolved in MeOH/AcOH 1:1 to a concentration of 1 mM. Then $NaBH_3CN$ (10 equiv.) was added and the mixture was stirred at rt in a glass vial for 3 h. This afforded **45** as a white solid in a 28% yield after purification using RP-HPLC.

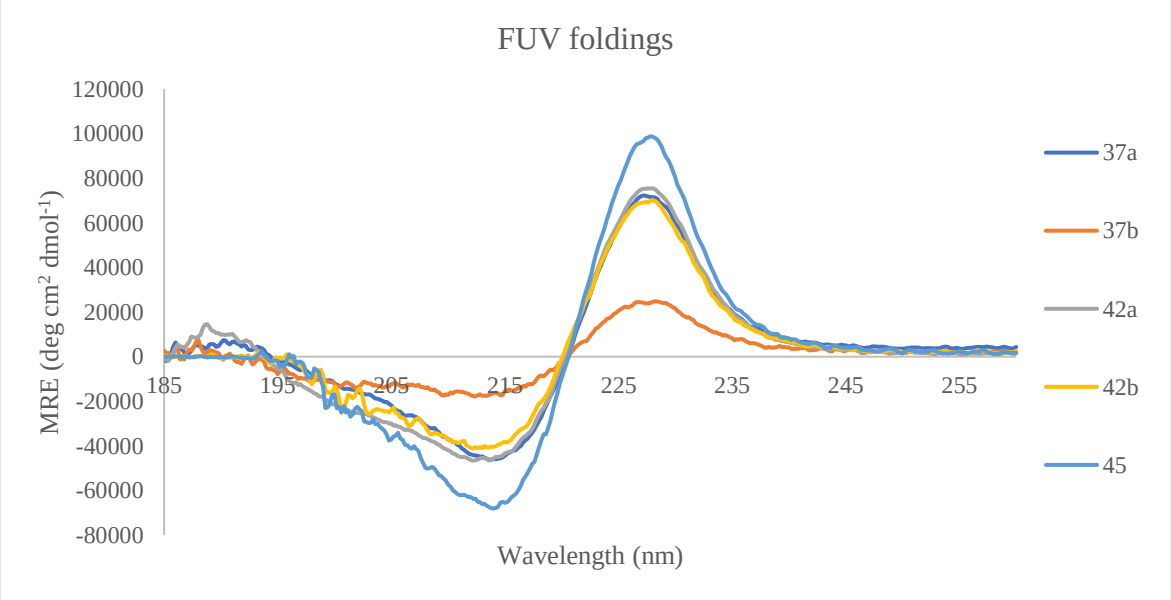
HRMS (ESI+) calcd for $C_{76}H_{103}N_{19}O_{17}$ $[M+2H]^{2+}$ 776.8884, found 776.8878. HPLC: 12.3 min 100% (20 min gradient), 24.6 min 98.6% (30 min gradient).

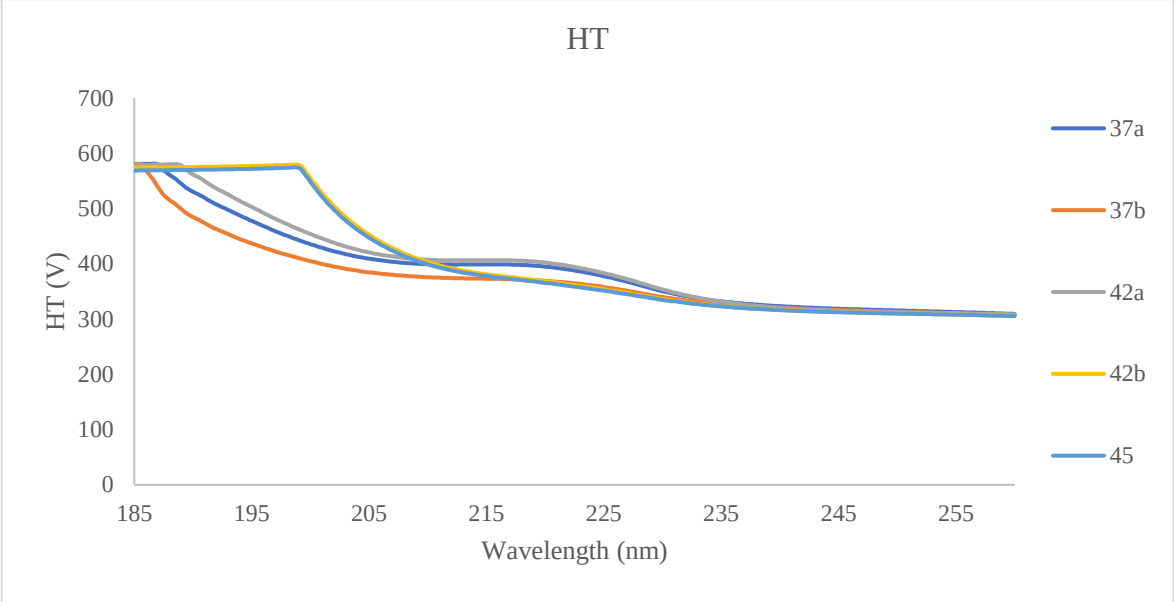
7.14.1 Circular dichroism

CD experiments were run at a 0.05 mg/mL concentration (30 µM) in sodium phosphate buffer (20 mM, pH 7) for **37a**, **37b** and **42a**. For **42b** and **45** a HEPES buffer (10 mM, pH 7) was used. A 0.1 cm cell was used for **42b** and **45**, and a 0.2 cm cell was used for **42a**, **37a** and **37b**.

Far UV experiments were run from 260 to 185 nm, at a 20 or 50 nm/min speed with a 2 sec response. Thermal denaturation experiments were performed in the far UV from 5 to 80 °C, monitoring every 5 °C. The CD was also monitored at 228 nm every 1 °C.

7.14.1.1 Far UV spectra





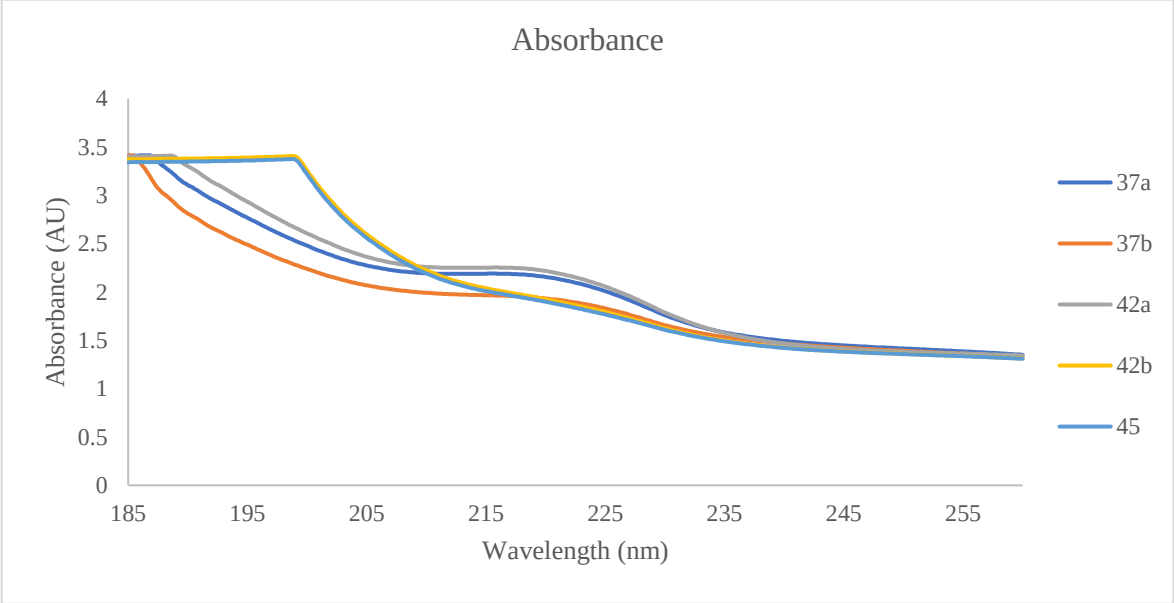


Figure 195. Far UV spectra for the peptides under study, showing the HT and absorbance graphs. Experiments for each peptide were run in the conditions described above.

7.14.1.2 Thermal denaturation

Monitoring in the far UV every 5 degrees, from 5 to 80 °C.

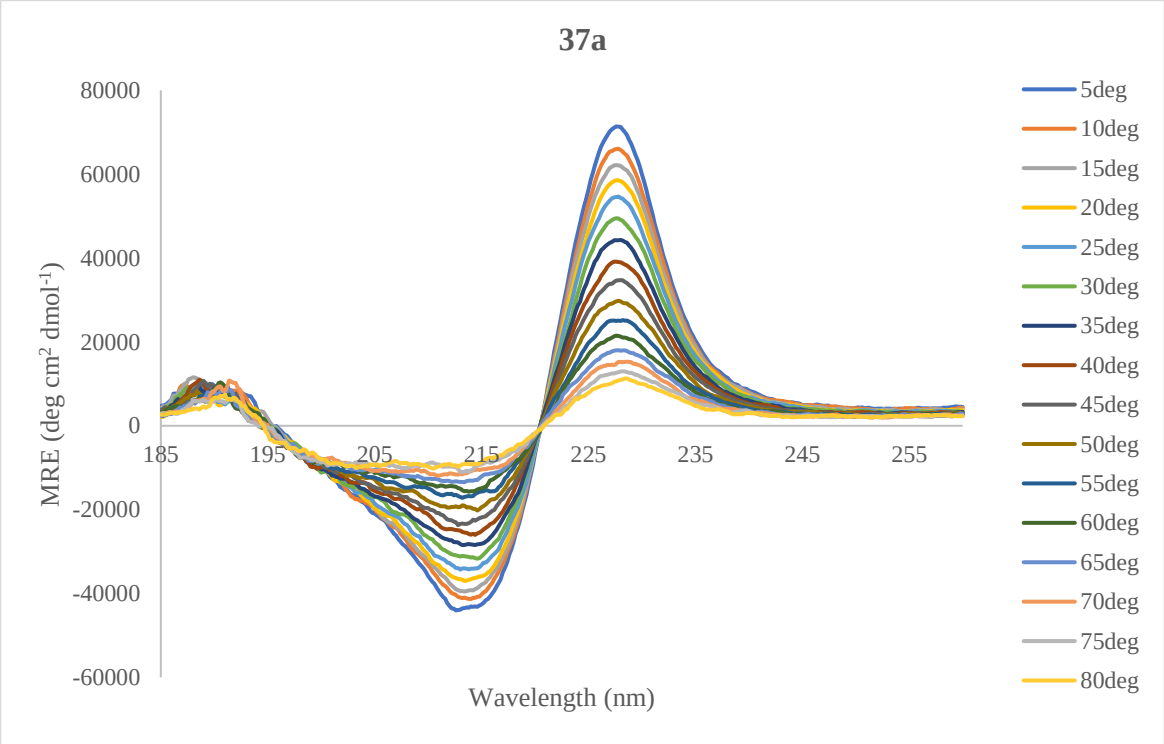


Figure 196. Thermal denaturation for 37a, run at a 0.05 mg/mL concentration in sodium phosphate buffer (20 mM, pH 7), in a 0.2 cm cell.

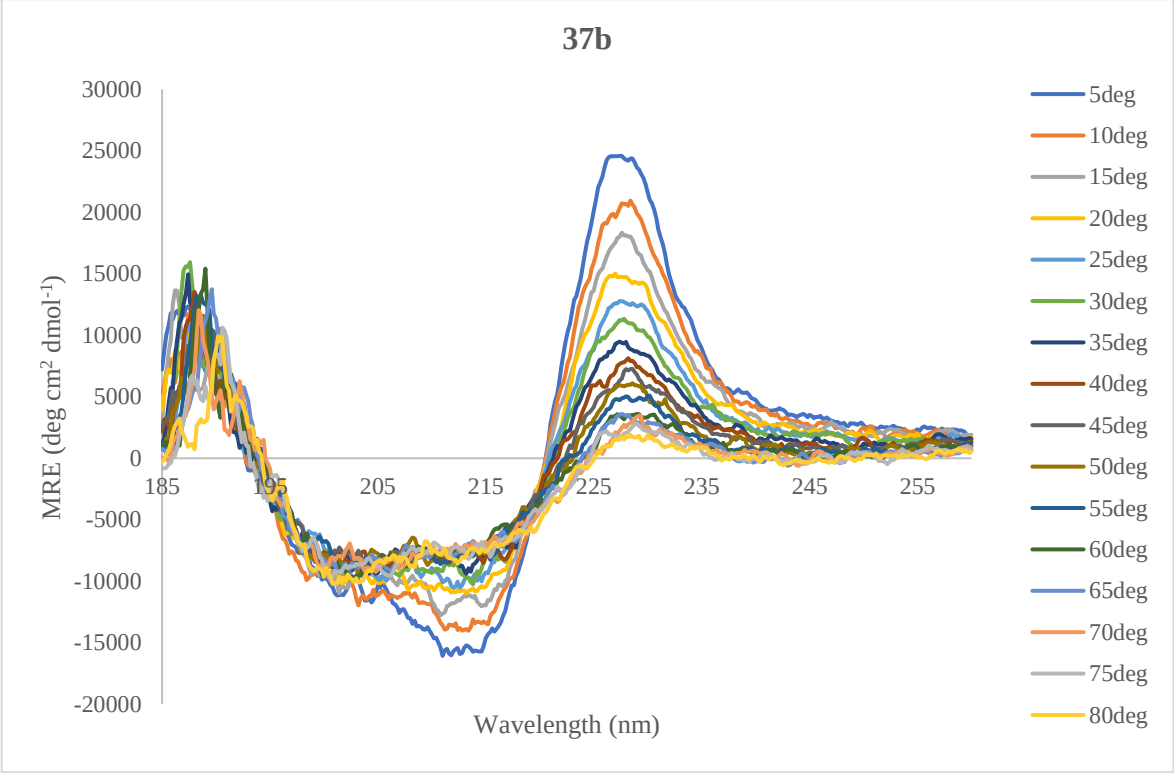


Figure 197. Thermal denaturation for 37b, run at a 0.05 mg/mL concentration in sodium phosphate buffer (20 mM, pH 7), in a 0.2 cm cell.

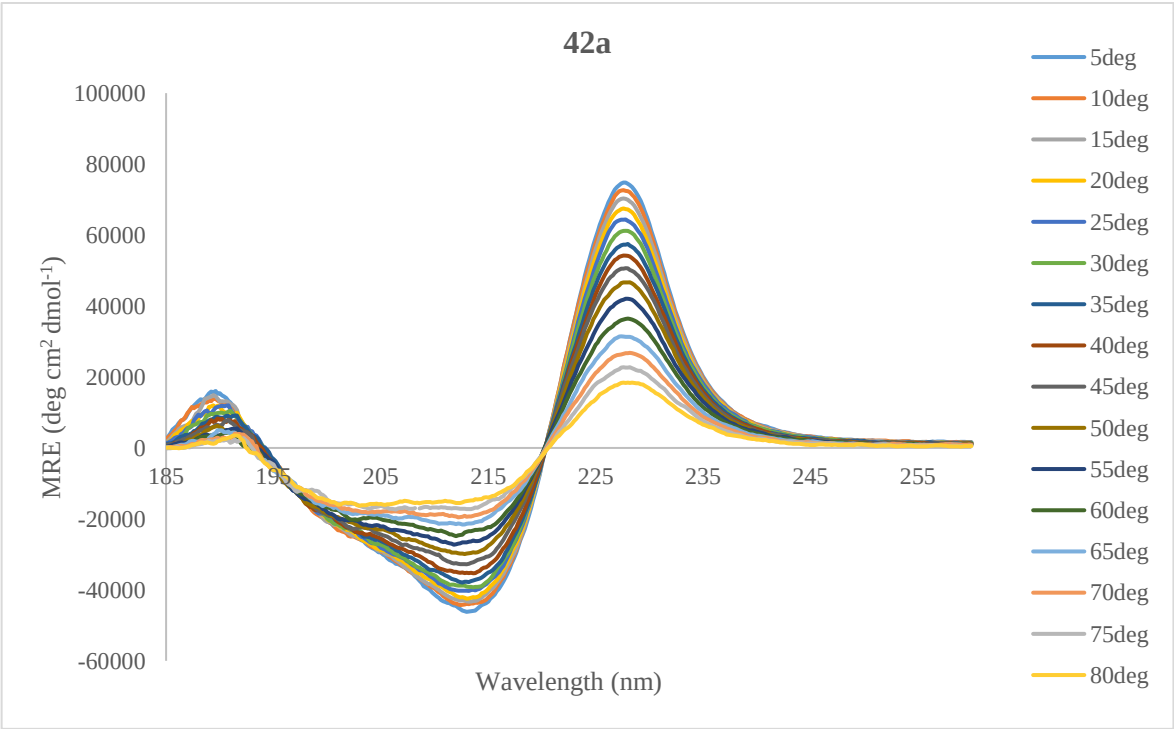


Figure 198. Thermal denaturation for 42a, run at a 0.05 mg/mL concentration in sodium phosphate buffer (20 mM, pH 7), in a 0.2 cm cell.

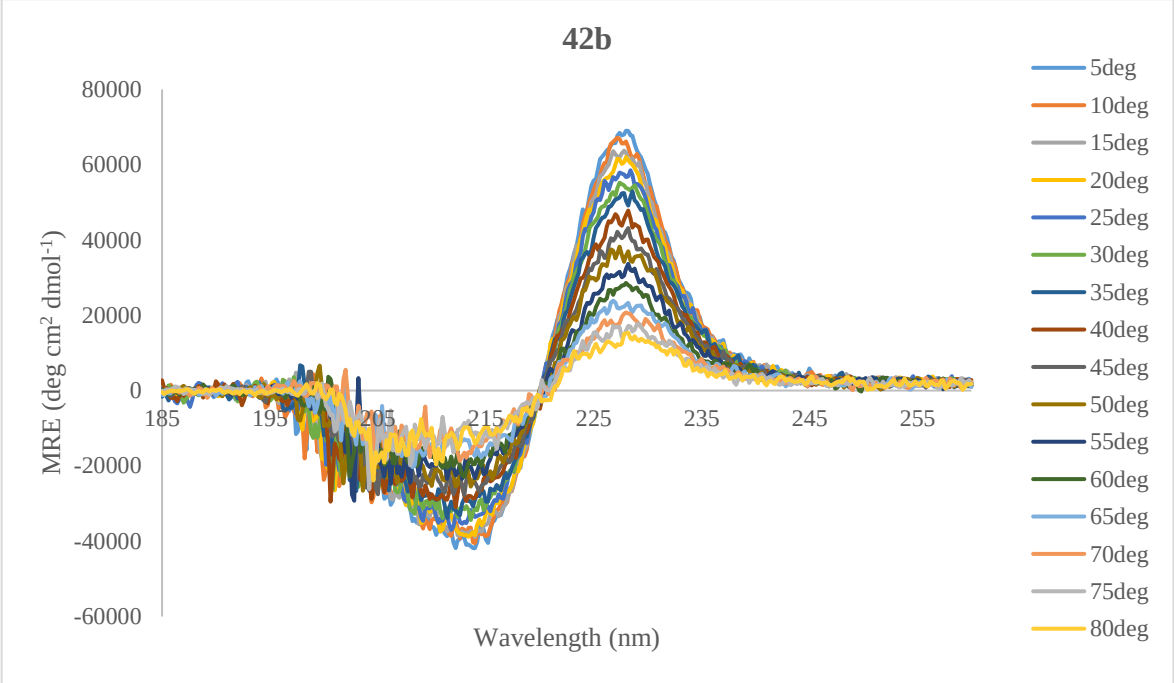


Figure 199. Thermal denaturation for 42b, run at a 0.05 mg/mL concentration in HEPES buffer (10 mM, pH 7), in a 0.1 cm cell.

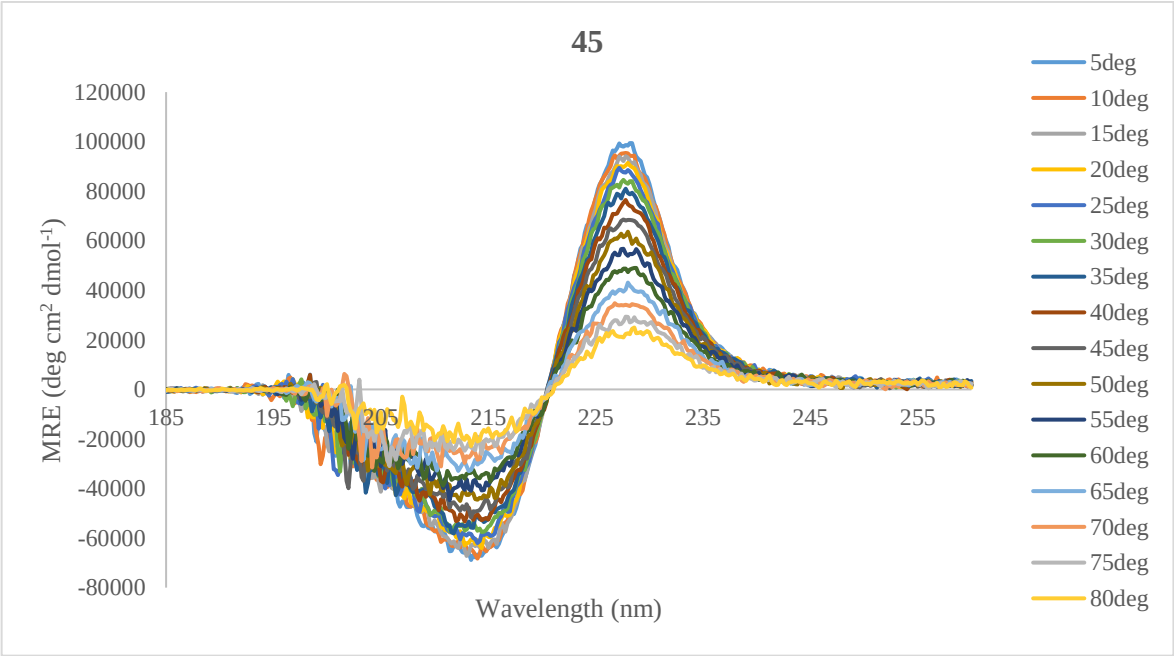
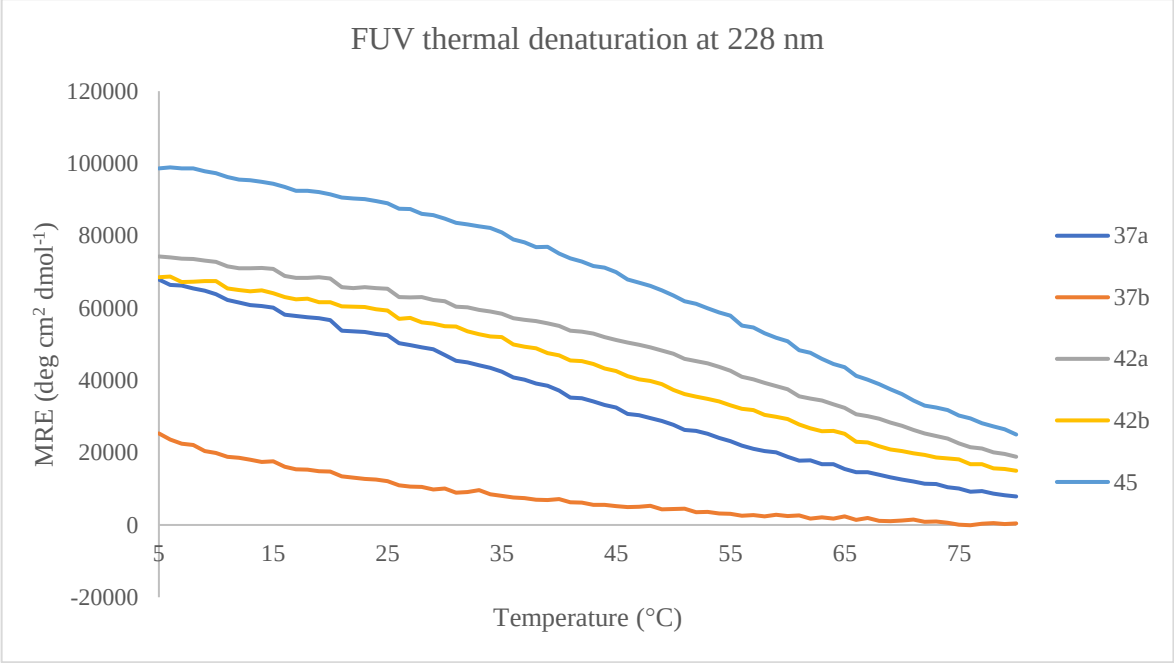
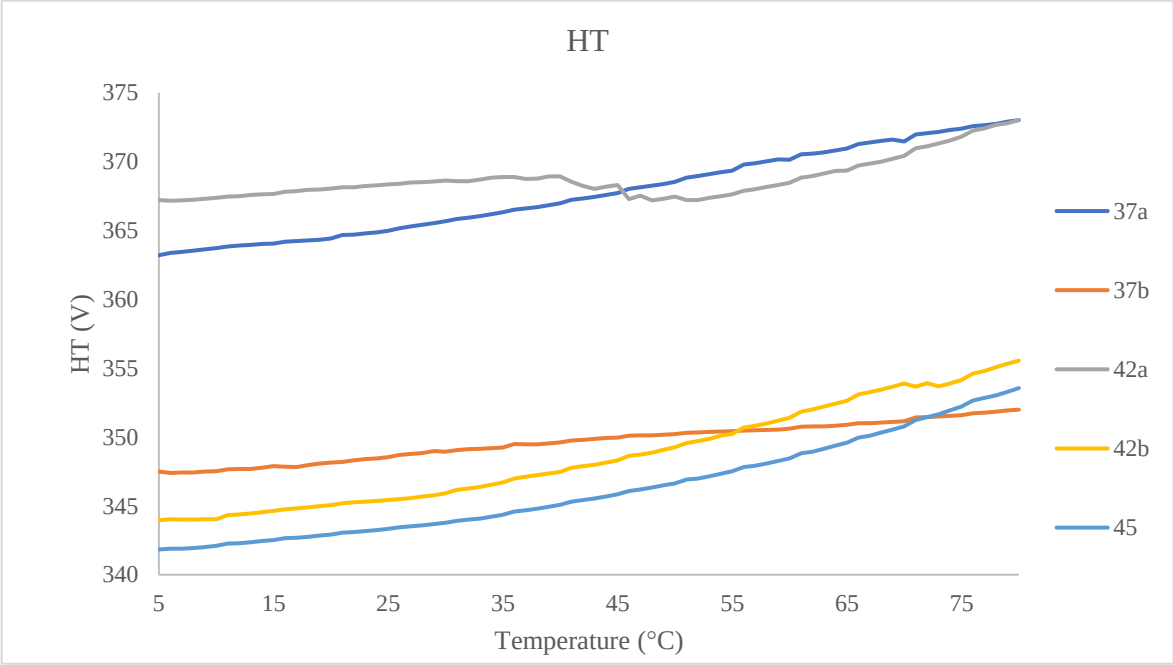


Figure 200. Thermal denaturation for 45, run at a 0.05 mg/mL concentration in HEPES buffer (10 mM, pH 7), in a 0.1 cm cell.


FUV thermal denaturation at 228 nm
120000
100000
80000
60000
40000
20000
0
-20000
MRE (deg cm2 dmol-1)
5
15
25
35
45
55
65
75
Temperature (°C)
37a
37b
42a
42b
45


HT
375
370
365
360
355
350
345
340
HT (V)
5
15
25
35
45
55
65
75
Temperature (°C)
37a
37b
42a
42b
45

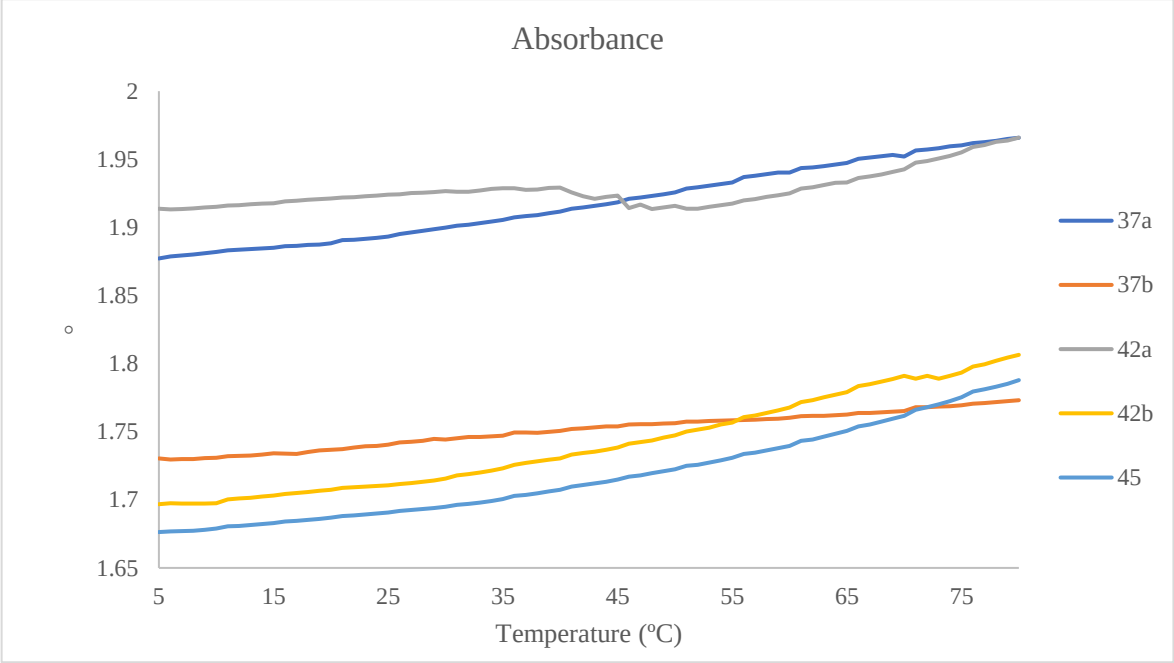


Figure 201. Thermal denaturation curves measured at 228 nm for all peptides under study, showing the HT and absorbance curves. Experiments for each peptide were run in the conditions described above.

8 References

(1) Markley, J. L.; Bax, A.; Arata, Y.; Hilbers, C. W.; Kaptein, R.; Sykes, B. D.; Wright, P. E.; Wüthrich, K. *J. Mol. NMR* **1998**, *12*, 1–23.

(2) Bosshard, H. R.; Marti, D. N.; Jelesarov, I. *J. Mol. Recognit.* **2004**, *17* (1), 1–16.

(3) Whitford, D. *Proteins: structure and function*, 1st Ed.; John Wiley and Sons: Chichester, UK, 2005.

(4) Creighton, T. E. *Biochem. J.* **1990**, *270*, 1–16.

(5) Pal, D.; Chakrabarti, P. *J. Mol. Biol.* **1999**, *294* (1), 271–288.

(6) Dugave, C.; Demange, L. *Chem. Rev.* **2003**, *103* (7), 2475–2532.

(7) Horovitz, A.; Fersht, A. R. *J. Mol. Biol.* **1990**, *214* (3), 613–617.

(8) Horovitz, A. *Fold. Des.* **1996**, *1* (6), R121–R126.

(9) Cockroft, S. L.; Hunter, C. A. *Chem. Soc. Rev.* **2007**, *36* (2), 172–188.

(10) Tatko, C. D.; Waters, M. L. *Protein Sci.* **2003**, *12* (11), 2443–2452.

(11) Tatko, C. D.; Waters, M. L. *J. Am. Chem. Soc.* **2004**, *126* (7), 2028–2034.

(12) Kuo, H. T.; Fang, C. J.; Tsai, H. Y.; Yang, M. F.; Chang, H. C.; Liu, S. L.; Kuo, L. H.; Wang, W. R.; Yang, P. A.; Huang, S. J.; Huang, S. L.; Cheng, R. P. *Biochemistry* **2013**, *52* (51), 9212–9222.

(13) Searle, M. S.; Griffiths-Jones, S. R.; Skinner-Smith, H. *J. Am. Chem. Soc.* **1999**, *121* (50), 11615–11620.

(14) Di Cera, E. *Chem. Rev.* **1998**, *98* (4), 1563–1591.

(15) Nelson, D. L.; Michael M Cox. *Lehninger Principles of Biochemistry*, 4th ed.; W H Freeman: New York, 2005.

(16) Montalbetti, C. A. G. N.; Falque, V. *Tetrahedron* **2005**, *61* (46), 10827–10852.

(17) El-Faham, A.; Albericio, F. *Chem. Rev.* **2011**, *111* (11), 6557–6602.

(18) Bruce Martin, R. *Biopolymers* **1998**, *45* (5), 351–353.

(19) Diez, M.; Petuya, V.; Urizar, M.; Macho, E.; Altuzarra, O. *New Trends in Mechanism and Machine Science*; Viadero, F., Ceccarelli, M., Eds.; Mechanisms and Machine Science; Springer Netherlands: Dordrecht, 2013; Vol. 7.

(20) Poole, L. B. *Free Radic. Biol. Med.* **2015**, *80*, 148–157.

(21) Thornton, J. M. *J. Mol. Biol.* **1981**, *151*, 261–287.

(22) Pijning, A. E.; Chiu, J.; Yeo, R. X.; Wong, J. W. H.; Hogg, P. J. *R. Soc. Open Sci.* **2018**, *5* (2).

(23) Craig, D. B.; Dombkowski, A. A. *BMC Bioinformatics* **2013**, *14*, 346.

(24) Li, L.; Mirny, L. A.; Shakhnovich, E. I. *Nat. Struct. Biol.* **2000**, *7* (4), 336–342.

(25) Clarke, J.; Fersht, A. R. *Biochemistry* **1993**, *32* (16), 4322–4329.

(26) Abkevich, V. I.; Shakhnovich, E. I. *J. Mol. Biol.* **2000**, *300* (4), 975–985.

(27) Markiewicz, B. N.; Yang, L.; Culik, R. M.; Gao, Y. Q.; Gai, F. *J. Phys. Chem. B* **2014**, *118* (12), 3317–3325.

(28) Tidor, B.; Karplus, M. *Proteins Struct. Funct. Bioinforma.* **1993**, *15* (1), 71–79.

(29) Pace, C. N.; Grimsley, G. R.; Thomson, J. A.; Barnett, B. J. *J. Biol. Chem.* **1988**, *263* (25), 11820–11825.

(30) Góngora-Benítez, M.; Tulla-Puche, J.; Albericio, F. *Chem. Rev.* **2014**, *114* (2), 901–926.

(31) Matsumura, M.; Becktel, W. J.; Levitt, M.; Matthews, B. W. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86* (17), 6562–6566.

(32) Narhi, L. O.; Arakawa, T.; Fox, G. M.; Tsai, L.; Rosenfeld, R.; Holst, P.; Miller, J. A.; Hua, Q. X.; Weiss, M. A. *Biochemistry* **1993**, *32* (19), 5214–5221.

(33) Giles, N. M.; Giles, G. I.; Jacob, C. *Biochem. Biophys. Res. Commun.* **2003**, *300* (1), 1–4.

(34) Jacob, C.; Giles, G. I.; Giles, N. M.; Sies, H. *Angew. Chem. - Int. Ed.* **2003**, *42* (39), 4742–4758.

(35) Schmidt, B.; Ho, L.; Hogg, P. J. *Biochemistry* **2006**, *45* (24), 7429–7433.

(36) Dombkowski, A. A.; Sultana, K. Z.; Craig, D. B. *FEBS Lett.* **2014**, *588* (2), 206–212.

(37) Wouters, M.; George, R.; Haworth, N. *Curr. Protein Pept. Sci.* **2007**, *8* (5), 484–495.

(38) Carugo, O.; Cemazar, M.; Zahariev, S.; Hudaky, I.; Gaspari, Z.; Perczel, A.; Pongor, S. *Protein Eng.* **2003**, *16* (9), 637–639.

(39) Richardson, J. S.; Videau, L. L.; Williams, C. J.; Richardson, D. C. *J. Mol. Biol.* **2017**, *429* (9), 1321–1335.

(40) Wang, X. H.; Connor, M.; Smith, R.; Maciejewski, M. W.; Howden, M. E. H.; Nicholson, G. M.; Christie, M. J.; King, G. F. *Nat. Struct. Biol.* **2000**, *7* (6), 505–513.

(41) Ruggles, E. L.; Deker, P. B.; Hondal, R. J. *Tetrahedron* **2009**, *65* (7), 1257–1267.

(42) Berg, J. M.; Tymoczko, J. L.; Stryer, L. In *Biochemistry*; W H Freeman: New York, 2002.

(43) Dill, K. A. *Biochemistry* **1990**, *29* (31), 7133–7155.

(44) Williamson, M. *How proteins work*; Garland Science: New York, 2012.

(45) Maccallum, P. H.; Poet, R.; Milner-White, E. J. *J. Mol. Biol.* **1995**, *248* (2), 374–384.

(46) Wang, L.; O'Connell, T.; Tropsha, A.; Hermans, J. *J. Mol. Biol.* **1996**, *262* (2), 283–293.

(47) Kumar, S.; Nussinov, R. *J. Mol. Biol.* **1999**, *293* (5), 1241–1255.

(48) Nick Pace, C.; Martin Scholtz, J.; Grimsley, G. R. *FEBS Lett.* **2014**, *588* (14), 2177–2184.

(49) Shi, Z.; Olson, C. A.; Bell, A. J.; Kallenbach, N. R. *Biopolym. - Pept. Sci. Sect.* **2001**, *60* (5), 366–380.

(50) Ciani, B.; Jourdan, M.; Searle, M. S. *J. Am. Chem. Soc.* **2003**, *125* (30), 9038–9047.

(51) Sippel, K. H.; Quiocho, F. A. *Protein Sci.* **2015**, *24* (7), 1040–1046.

(52) Sticke, D. F.; Presta, L. G.; Dill, K. A.; Rose, G. D. *J. Mol. Biol.* **1992**, *226* (4), 1143–1159.

(53) Schulz, G. E.; Schirmer, R. H. *Principles of protein structure*; Cantor, C. R., Ed.; Springer: New York, 1979.

(54) Van der Lubbe, S. C. C.; Fonseca Guerra, C. *Chem. - An Asian J.* **2019**, *14* (16), 2760–2769.

(55) Herschlag, D.; Pinney, M. M. *Biochemistry* **2018**, *57* (24), 3338–3352.

(56) Mo, Y.; Wang, C.; Guan, L.; Braïda, B.; Hiberty, P. C.; Wu, W. *Chem. - A Eur. J.* **2014**, *20* (27), 8444–8452.

(57) Newberry, R. W.; Raines, R. T. *Acc. Chem. Res.* **2017**, *50* (8), 1838–1846.

(58) Newberry, R. W.; Vanveller, B.; Guzei, I. A.; Raines, R. T. *J. Am. Chem. Soc.* **2013**, *135* (21), 7843–7846.

(59) Bartlett, G. J.; Choudhary, A.; Raines, R. T.; Woolfson, D. N. *Nat. Chem. Biol.* **2010**, *6* (8), 615–620.

(60) Pandey, A. K.; Naduthambi, D.; Thomas, K. M.; Zondlo, N. J. *J. Am. Chem. Soc.* **2013**, *135* (11), 4333–4363.

(61) Quiñonero, D.; Garau, C.; Rotger, C.; Frontera, A.; Ballester, P.; Costa, A.; Deyà, P. M. *Angew. Chem. Int. Ed.* **2002**, *41* (18), 3389–3392.

(62) Lucas, X.; Bauzá, A.; Frontera, A.; Quiñonero, D. *Chem. Sci.* **2016**, *7* (2), 1038–1050.

(63) Smith, M. S.; Lawrence, E. E. K.; Billings, W. M.; Larsen, K. S.; Bécar, N. A.; Price, J. L. *ACS Chem. Biol.* **2017**, *12* (10), 2535–2537.

(64) Gallivan, J. P.; Dougherty, D. A. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (17), 9459–9464.

(65) Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, *97* (5), 1303–1324.

(66) Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem. - Int. Ed.* **2003**, *42* (11), 1210–1250.

(67) Nishio, M.; Umezawa, Y.; Fantini, J.; Weiss, M. S.; Chakrabarti, P. *Phys. Chem. Chem. Phys.* **2014**, *16* (25), 12648–12683.

(68) Tóth, G.; Murphy, R. F.; Lovas, S. *Protein Eng.* **2001**, *14* (8), 543–547.

(69) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2000**, *122* (15), 3746–3753.

(70) Brandl, M.; Weiss, M. S.; Jabs, A.; Sühnel, J.; Hilgenfeld, R. *J. Mol. Biol.* **2001**, *307* (1), 357–377.

(71) Zondlo, N. J. *Acc. Chem. Res.* **2013**, *46* (4), 1039–1049.

(72) Crothers, D. M.; Zimm, B. H. *J. Mol. Biol.* **1964**, *9* (1), 1–9.

(73) Burley, S. K.; Petsko, G. A. *Science* **1985**, *229* (4708), 23–28.

(74) Yang, A. S.; Honig, B. *J. Mol. Biol.* **1995**, *252* (3), 366–376.

(75) Waters, M. L. *Biopolym. - Pept. Sci. Sect.* **2004**, *76* (5), 435–445.

(76) Carver, F. J.; Hunter, C. A.; Seward, E. M. *Chem. Commun.* **1998**, No. 2, 775–776.

(77) Hunter, C. A.; Singh, J.; Thornton, J. M. *J. Mol. Biol.* **1991**, *218* (4), 837–846.

(78) Hunter, C. A.; Sanders, J. K. M. *J. Am. Chem. Soc.* **1990**, *112* (14), 5525–5534.

(79) Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J. *J. Chem. Soc. Perkin Trans. 2* **2001**, No. 5, 651–669.

(80) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci.* **2001**, *98* (10), 5578–5583.

(81) Rodriguez-Larrea, D.; Ibarra-Molero, B.; Sanchez-Ruiz, J. M. *Biophys. J.* **2006**, *91* (5), L48–L50.

(82) Cheung, M. S.; García, A. E.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (2), 685–690.

(83) Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1994**, *3* (2), 211–226.

(84) Baldwin, R. L. *J. Mol. Biol.* **2007**, *371* (2), 283–301.

(85) Fernández, A.; Sosnick, T. R.; Colubri, A. *J. Mol. Biol.* **2002**, *321* (4), 659–675.

(86) Hunter, C. A. *Angew. Chem. - Int. Ed.* **2004**, *43* (40), 5310–5324.

(87) Misra, G. *Introduction to biomolecular structure and biophysics: Basics of biophysics*; Springer: Singapore, 2017.

(88) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7* (1), 95–99.

(89) Lovell, S. C.; Davis, I. W.; Arendall, W. B.; Bakker, P. I. W. de; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. *Proteins Struct. Funct. Bioinforma.* **2003**, *50* (3), 437–450.

(90) Williams, C. J.; Headd, J. J.; Moriarty, N. W.; Prisant, M. G.; Videau, L. L.; Deis, L. N.; Verma, V.; Keedy, D. A.; Hintze, B. J.; Chen, V. B.; Jain, S.; Lewis, S. M.; Arendall, W. B.; Snoeyink, J.; Adams, P. D.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *Protein Sci.* **2018**, *27* (1), 293–315.

(91) Fischer, G. *Chem. Soc. Rev.* **2000**, *29* (2), 119–127.

(92) Weiss, M. S.; Jabs, A.; Hilgenfeld, R. *Nat. Struct. Mol. Biol.* **1998**, *5* (8), 676.

(93) Azzarito, V.; Long, K.; Murphy, N. S.; Wilson, A. J. *Nat. Chem.* **2013**, *5* (3), 161–173.

(94) Jayatunga, M. K. P.; Thompson, S.; Hamilton, A. D. *Bioorganic Med. Chem. Lett.* **2014**, *24* (3), 717–724.

(95) Holland, G. P.; Creager, M. S.; Jenkins, J. E.; Lewis, R. V.; Yarger, J. L. *J. Am. Chem. Soc.* **2008**, *130* (30), 9871–9877.

(96) Parkhe, A. D.; Seeley, S. K.; Gardner, K.; Thompson, L.; Lewis, R. V. *J. Mol. Recognit.* **1997**, *10* (1), 1–6.

(97) Loughlin, W. A.; Tyndall, J. D. A.; Glenn, M. P.; Hill, T. A.; Fairlie, D. P. *Chem. Rev.* **2010**, *110* (6), PR32–PR69.

(98) Somers, W. S.; Phillips, S. E. V. *Nature* **1992**, *359* (6394), 387–393.

(99) Watkins, A. M.; Arora, P. S. *ACS Chem. Biol.* **2014**, *9* (8), 1747–1754.

(100) Ruiz-Gómez, G.; Tyndall, J. D. A.; Pfeiffer, B.; Abbenante, G.; Fairlie, D. P. *Chem. Rev.* **2005**, *105*, 793–826.

(101) Franklin, M. W.; Slusky, J. S. G. *J. Mol. Biol.* **2018**, *430* (18), 3251–3265.

(102) Sewald, N.; Jakubke, H. *Peptides: Chemistry and Biology*, 2nd Ed.; Wiley-VCH: Weinheim, 2009.

(103) Kuo, N. N. W.; Huang, J. J. T.; Miksovska, J.; Chen, R. P. Y.; Larsen, R. W.; Chan, S. I. *J. Am. Chem. Soc.* **2005**, *127* (48), 16945–16954.

(104) Sharpe, T.; Jonsson, A. L.; Rutherford, T. J.; Daggett, V.; Fersht, A. R. *Protein Sci.* **2007**, *16* (10), 2233–2239.

(105) Griffiths-Jones, S. R.; Maynard, A. J.; Sharman, G. J.; Searle, M. S. *Chem. Commun.* **1998**, *1* (7), 789–790.

(106) Gupta, S.; Sasidhar, Y. U. *J. Phys. Chem. B* **2017**, *121* (6), 1268–1283.

(107) Lahiri, P.; Verma, H.; Ravikumar, A.; Chatterjee, J. *Chem. Sci.* **2018**, *9* (20), 4600–4609.

(108) Marcelino, A. M. C.; Gierasch, L. M. *Biopolymers* **2008**, *89* (5), 380–391.

(109) Chou, K. C. *Anal. Biochem.* **2000**, *286* (1), 1–16.

(110) Hutchinson, E. G.; Thornton, J. M. *Protein Sci.* **1994**, *3* (12), 2207–2216.

(111) De Brevern, A. G. *Sci. Rep.* **2016**, *6*, 1–15.

(112) Rajashankar, K. R.; Ramakumar, S. *Protein Sci.* **1996**, *5* (5), 932–946.

(113) Venkatachalam, C. M. *Biopolymers* **1968**, *6* (10), 1425–1436.

(114) McCallister, E. L.; Alm, E.; Baker, D. *Nat. Struct. Biol.* **2000**, *7* (8), 669–673.

(115) Byrne, C.; Belnou, M.; Baulieu, E. E.; Lequin, O.; Jacquot, Y. *Pept. Sci.* **2019**, *111* (4), 1–14.

(116) Wedemeyer, W. J.; Welker, E.; Scheraga, H. A. *Biochemistry* **2002**, *41* (50), 14637–14644.

(117) Schmidpeter, P. A. M.; Schmid, F. X. *J. Mol. Biol.* **2015**, *427* (7), 1609–1631.

(118) Andreotti, A. H. *Biochemistry* **2003**, *42* (32), 9515–9524.

(119) McNulty, J. C.; Jackson, P. J.; Thompson, D. A.; Chai, B.; Gantz, I.; Barsh, G. S.; Dawson, P. E.; Millhauser, G. L. *J. Mol. Biol.* **2005**, *346* (4), 1059–1070.

(120) Keller, M.; Boissard, C.; Patiny, L.; Chung, N. N.; Lemieux, C.; Mutter, M.; Schiller, P. W. *J. Med. Chem.* **2001**, *44* (23), 3896–3903.

(121) Searle, M. S. In *Protein Folding Handbook*; Buchner, J., Kiefhaber, T., Eds.;

WILEY-VCH: Weinheim, 2005; Vol. 1, pp 314–342.

(122) Haque, T. S.; Gellman, S. H. *J. Am. Chem. Soc.* **1997**, *119* (9), 2303–2304.

(123) Stanger, H. E.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120* (17), 4236–4237.

(124) Wu, L.; McElheny, D.; Huang, R.; Keiderling, T. A. *Biochemistry* **2009**, *48* (43), 10362–10371.

(125) Tatko, C. D.; Waters, M. L. *J. Am. Chem. Soc.* **2002**, *124* (32), 9372–9373.

(126) Santiveri, C. M.; León, E.; Rico, M.; Jiménez, M. A. *Chem. - A Eur. J.* **2008**, *14* (2), 488–499.

(127) Robinson, J. A.; Demarco, S.; Gombert, F.; Moehle, K.; Obrecht, D. *Drug Discov. Today* **2008**, *13* (21–22), 944–951.

(128) Wong, H. C.; Mao, J.; Nguyen, J. T.; Srinivas, S.; Zhang, W.; Liu, B.; Li, L.; Wu, D.; Zheng, J. *Nat. Struct. Biol.* **2000**, *7* (12), 1178–1184.

(129) Hillier, B. J.; Christopherson, K. S.; Prehoda, K. E.; Bredt, D. S.; Lim, W. A. *Science* **1999**, *284* (5415), 812–815.

(130) Gajiwala, K. S.; Chen, H.; Cornille, F.; Roques, B. P.; Reith, W.; Mach, B.; Burley, S. K. *Nature* **2000**, *403* (6772), 916–921.

(131) Zavala-Ruiz, Z.; Strug, I.; Walker, B. D.; Norris, P. J.; Stern, L. J. *Proc. Natl. Acad. Sci.* **2004**, *101* (36), 13279–13284.

(132) Sivanesam, K.; Kier, B. L.; Whedon, S. D.; Chatterjee, C.; Andersen, N. H. *FEBS Lett.* **2016**, *590* (24), 4480–4488.

(133) Xu, L.; Chou, S.; Wang, J.; Shao, C.; Li, W.; Zhu, X.; Shan, A. *Amino Acids* **2015**, *47* (11), 2385–2397.

(134) Podewin, T.; Rampp, M. S.; Turkanovic, I.; Karaghiosoff, K. L.; Zinth, W.; Hoffmann-Röder, A. *Chem. Commun.* **2015**, *51* (19), 4001–4004.

(135) Dong, S. L.; Löweneck, M.; Schrader, T. E.; Schreier, W. J.; Zinth, W.; Moroder, L.; Renner, C. *Chem. - A Eur. J.* **2006**, *12* (4), 1114–1120.

(136) Pelay-Gimeno, M.; Glas, A.; Koch, O.; Grossmann, T. N. *Angew. Chem. - Int. Ed.* **2015**, *54* (31), 8896–8927.

(137) Davis, J. M.; Tsou, L. K.; Hamilton, A. D. *Chem. Soc. Rev.* **2007**, *36* (2), 326–334.

(138) Saraogi, I.; Hebda, J. A.; Becerril, J.; Estroff, L. A.; Miranker, A. D.; Hamilton, A. D. *Angew. Chem. - Int. Ed.* **2010**, *49* (4), 736–739.

(139) Lau, Y. H.; de Andrade, P.; Wu, Y.; Spring, D. R. *Chem. Soc. Rev.* **2015**, *44* (1), 91–102.

(140) Patgiri, A.; Jochim, A. L.; Arora, P. S. *Acc. Chem. Res.* **2008**, *41* (10), 1289–1300.

(141) Morales, P.; Jiménez, M. A. *Arch. Biochem. Biophys.* **2019**, *661* (2019), 149–167.

(142) Cabezas, E.; Satterthwait, A. C. *J. Am. Chem. Soc.* **1999**, *121* (16), 3862–3875.

(143) Sawyer, N.; Arora, P. S. *ACS Chem. Biol.* **2018**, *13* (8), 2027–2032.

(144) Choudhary, A.; Raines, R. T. *ChemBioChem* **2011**, *12* (12), 1801–1807.

(145) Tam, A.; Arnold, U.; Soellner, M. B.; Raines, R. T. *J. Am. Chem. Soc.* **2007**, *129* (42), 12670–12671.

(146) Vanderesse, R.; Thevenet, L.; Marraud, M.; Boggetto, N.; Reboud, M.; Corbier, C. *J. Pept. Sci.* **2003**, *9* (5), 282–299.

(147) Miller, S. J.; Grubbs, R. H. *J. Am. Chem. Soc.* **1995**, *117* (21), 5855–5856.

(148) Vartak, A. P.; Skoblenick, K.; Thomas, N.; Mishra, R. K.; Johnson, R. L. *J. Med. Chem.* **2007**, *50* (26), 6725–6729.

(149) Merritt, H. I.; Sawyer, N.; Arora, P. S. *Pept. Sci.* **2020**, No. 24145, 1–22.

(150) Deplazes, E.; Chin, Y. K. Y.; King, G. F.; Mancera, R. L. *Proteins Struct. Funct. Bioinforma.* **2020**, *88* (3), 485–502.

(151) Cochran, A. G.; Tong, R. T.; Starovasnik, M. A.; Park, E. J.; McDowell, R. S.; Theaker, J. E.; Skelton, N. J. *J. Am. Chem. Soc.* **2001**, *123* (4), 625–632.

(152) Park, J. H.; Waters, M. L. *Org. Biomol Chem.* **2013**, *11* (1), 69–77.

(153) Diana, D.; Di Salvo, C.; Celentano, V.; De Rosa, L.; Romanelli, A.; Fattorusso, R.; D'Andrea, L. D. *Org. Biomol. Chem.* **2018**, *16* (5), 787–795.

(154) Celentano, V.; Diana, D.; De Rosa, L.; Romanelli, A.; Fattorusso, R.; D'Andrea, L. D. *Chem. Commun.* **2012**, *48* (5), 762–764.

(155) Holland-Nell, K.; Meldal, M. *Angew. Chem. - Int. Ed.* **2011**, *50* (22), 5204–5206.

(156) Nair, R. V.; Baravkar, S. B.; Ingole, T. S.; Sanjayan, G. J. *Chem. Commun.* **2014**, *50* (90), 13874–13884.

(157) Khakshoor, O.; Nowick, J. S. *Curr. Opin. Chem. Biol.* **2008**, *12* (6), 722–729.

(158) Venkatraman, J.; Shankaramma, S. C.; Balaram, P. *Chem. Rev.* **2001**, *101* (10), 3131–3152.

(159) Shankaramma, S. C.; Athanassiou, Z.; Zerbe, O.; Moehle, K.; Mouton, C.; Bernardini, F.; Vrijbloed, J. W.; Obrecht, D.; Robinson, J. A. *ChemBioChem* **2002**, *3* (11), 1126–1133.

(160) Fasan, R.; Dias, R. L. A.; Moehle, K.; Zerbe, O.; Vrijbloed, J. W.; Obrecht, D.; Robinson, J. A. *Angew. Chem. - Int. Ed.* **2004**, *43* (16), 2109–2112.

(161) Nair, C. M.; Vijayan, M.; Venkatachalapathi, Y. V.; Balaram, P. *J. Chem. Soc., Chem. Commun.* **1979**, No. 24, 1183–1184.

(162) Robinson, J. A. *Acc. Chem. Res.* **2008**, *41* (10), 1278–1288.

(163) Nowick, J. S.; Brower, J. O. *J. Am. Chem. Soc.* **2003**, *125* (4), 876–877.

(164) Cameron, A. J.; Varnava, K. G.; Edwards, P. J. B.; Harjes, E.; Sarojini, V. *J. Pept. Sci.* **2018**, *24* (3094), 1–10.

(165) Woods, R. J.; Brower, J. O.; Castellanos, E.; Hashemzadeh, M.; Khakshoor, O.; Russu, W. A.; Nowick, J. S. *J. Am. Chem. Soc.* **2007**, *129* (9), 2548–2558.

(166) Aravinda, S.; Shamala, N.; Rajkishore, R.; Gopi, H. N.; Balaram, P. *Angew. Chem. - Int. Ed.* **2002**, *41* (20), 3863–3865.

(167) Masterson, L. R.; Etienne, M. A.; Porcelli, F.; Barany, G.; Hammer, R. P.; Veglia,

G. *Pept. Sci.* **2007**, *88* (5), 746–753.

(168) Raghavender, U. S.; Aravinda, S.; Rai, R.; Shamala, N.; Balaram, P. *Org. Biomol. Chem.* **2010**, *8* (14), 3133–3135.

(169) Oh, K.; Guan, Z. *Chem. Commun.* **2006**, *3* (29), 3069–3071.

(170) Doan, N. D.; Lubell, W. D. *Biopolymers* **2015**, *104* (5), 629–635.

(171) Chingle, R.; Proulx, C.; Lubell, W. D. *Acc. Chem. Res.* **2017**, *50* (7), 1541–1556.

(172) Fink, B. E.; Kym, P. R.; Katzenellenbogen, J. A. *J. Am. Chem. Soc.* **1998**, *120* (18), 4334–4344.

(173) Nagai, U.; Sato, K. *Tetrahedron Lett.* **1985**, *26* (5), 647–650.

(174) Eckhardt, B.; Grosse, W.; Essen, L. O.; Geyer, A. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (43), 18336–18341.

(175) Godina, T. A.; Lubell, W. D. *J. Org. Chem.* **2011**, *76* (14), 5846–5849.

(176) Nowick, J. S. *Acc. Chem. Res.* **2008**, *41* (10), 1319–1330.

(177) Gokhale, A.; Weldeghiorghis, T. K.; Taneja, V.; Satyanarayanajois, S. D. *J. Med. Chem.* **2011**, *54* (15), 5307–5319.

(178) Jha, A.; Kumar, M. G.; Gopi, H. N.; Paknikar, K. M. *Langmuir* **2018**, *34* (4), 1591–1600.

(179) Davidson, A.; Leeper, T. C.; Athanassiou, Z.; Patora-Komisarska, K.; Karn, J.; Robinson, J. A.; Varani, G. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (29), 11931–11936.

(180) Moehle, K.; Athanassiou, Z.; Patora, K.; Davidson, A.; Varani, G.; Robinson, J. A. *Angew. Chem. - Int. Ed.* **2007**, *46* (47), 9101–9104.

(181) Morse, R. P.; Willett, J. L. E.; Johnson, P. M.; Zheng, J.; Credali, A.; Iniguez, A.; Nowick, J. S.; Hayes, C. S.; Goulding, C. W. *J. Mol. Biol.* **2015**, *427* (23), 3766–3784.

(182) Rosenström, U.; Sköld, C.; Lindeberg, G.; Botros, M.; Nyberg, F.; Karlén, A.; Hallberg, A. *J. Med. Chem.* **2006**, *49* (20), 6133–6137.

(183) Gunnoo, S. B.; Madder, A. *Org. Biomol. Chem.* **2016**, *14* (34), 8002–8013.

(184) Jeet Kalia; Raines, R. *Curr Org Chem.* **2010**, *14* (2), 138–147.

(185) Lim, S. I.; Kwon, I. *Crit. Rev. Biotechnol.* **2016**, *36* (5), 803–815.

(186) Koniev, O.; Wagner, A. *Chem. Soc. Rev.* **2015**, *44* (15), 5495–5551.

(187) Spicer, C. D.; Pashuck, E. T.; Stevens, M. M. *Chem. Rev.* **2018**, *118* (16), 7702–7743.

(188) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science (80-.).* **1994**, *266* (5186), 776–779.

(189) Conibear, A. C.; Watson, E. E.; Payne, R. J.; Becker, C. F. W. *Chem. Soc. Rev.* **2018**, *47* (24), 9046–9068.

(190) Kulkarni, S. S.; Sayers, J.; Premdjee, B.; Payne, R. J. *Nat. Rev. Chem.* **2020**, *2* (4), 1–17.

(191) Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoët, M.; Monbaliu, J. C. M.; Melnyk, O. *Chem. Rev.* **2019**, *119* (12), 7328–7443.

(192) Rose, K. *J. Am. Chem. Soc.* **1994**, *116* (1), 30–33.

(193) Jiao, T.; Wu, G.; Zhang, Y.; Shen, L.; Lei, Y.; Wang, C.-Y.; Fahrenbach, A. C.; Li, H. *Angew. Chem. Int. Ed.* **2020**, *59*, 2–20.

(194) Kölmel, D. K.; Kool, E. T. *Chem. Rev.* **2017**, *117* (15), 10358–10376.

(195) Dirksen, A.; Hackeng, T. M.; Dawson, P. E. *Angew. Chem. - Int. Ed.* **2006**, *45* (45), 7581–7584.

(196) Kalia, J.; Raines, R. T. *Angew Chem Int Ed Engl.* **2008**, *47* (39), 7523–7526.

(197) Avrutina, O.; Schmoldt, H. U.; Gabrijelcic-Geiger, D.; Wentzel, A.; Frauendorf, H.; Sommerhoff, C. P.; Diederichsen, U.; Kolmar, H. *ChemBioChem* **2008**, *9* (1), 33–37.

(198) Chow, H. Y.; Zhang, Y.; Matheson, E.; Li, X. *Chem. Rev.* **2019**, *119* (17), 9971–10001.

(199) Torrini, I.; Nalli, M.; Paglialunga Paradisi, M.; Pagani Zecchini, G.; Lucente, G.; Spisani, S. *J. Pept. Res.* **2001**, *58* (1), 56–66.

(200) Hentzen, N. B.; Smeenk, L. E. J.; Witek, J.; Riniker, S.; Wennemers, H. *J. Am. Chem. Soc.* **2017**, *139* (36), 12815–12820.

(201) Haney, C. M.; Horne, W. S. *Chem. - A Eur. J.* **2013**, *19* (34), 11342–11351.

(202) Haney, C. M.; Loch, M. T.; Horne, W. S. *Chem. Commun.* **2011**, *47* (39), 10915–10917.

(203) Schmid, F. X.; Baldwin, R. L. *Proc. Natl Acad. Sci. USA* **1978**, *75* (10), 4764–4768.

(204) Vogel, M.; Bukau, B.; Mayer, M. P. *Mol. Cell* **2006**, *21* (3), 359–367.

(205) Nicholson, L. K.; Lu, K. P.; Finn, G.; Ho Lee, T. *Nat. Chem. Biol.* **2007**, *3* (10), 619–629.

(206) Reimer, U.; Fischer, G. *Biophys. Chem.* **2002**, *96* (2–3), 203–212.

(207) Sarkar, P.; Reichman, C.; Saleh, T.; Birge, R. B.; Kalodimos, C. G. *Mol Cell.* **2007**, *25* (3), 413–426.

(208) S.P.Sansom, M.; Weinstein, H. *Trends Pharmacol. Sci.* **2000**, *21* (11), 445–451.

(209) Jabs, A.; Weiss, M. S.; Hilgenfeld, R. *J. Mol. Biol.* **1999**, *286* (1), 291–304.

(210) Reimer, U.; Scherer, G.; Drewello, M.; Kruber, S.; Schutkowski, M.; Fischer, G. *J. Mol. Biol.* **1998**, *279* (2), 449–460.

(211) Wu, W. J.; Raleigh, D. P. *Biopolymers* **1998**, *45* (5), 381–394.

(212) Bhattacharyya, R.; Chakrabarti, P. *J. Mol. Biol.* **2003**, *331* (4), 925–940.

(213) Meng, H. Y.; Thomas, K. M.; Lee, A. E.; Zondlo, N. J. *Biopolym. (Peptide Sci.* **2006**, *84*, 192–204.

(214) Dasgupta, B.; Chakrabarti, P.; Basu, G. *FEBS Lett.* **2007**, *581* (23), 4529–4532.

(215) Ganguly, H. K.; Majumder, B.; Chattopadhyay, S.; Chakrabarti, P.; Basu, G. *J. Am.*

Chem. Soc. **2012**, *134* (10), 4661–4669.

(216) Chatterjee, J.; Rechenmacher, F.; Kessler, H. *Angew. Chem. - Int. Ed.* **2013**, *52* (1), 254–269.

(217) Räder, A. F. B.; Reichart, F.; Weinmüller, M.; Kessler, H. *Bioorganic Med. Chem.* **2018**, *26* (10), 2766–2773.

(218) Hilimire, T. A.; Bennett, R. P.; Stewart, R. A.; Garcia-Miranda, P.; Blume, A.; Becker, J.; Sherer, N.; Helms, E. D.; Butcher, S. E.; Smith, H. C.; Miller, B. L. *ACS Chem. Biol.* **2016**, *11* (1), 88–94.

(219) Chatterjee, J.; Gilon, C.; Hoffman, A.; Kessler, H. *Acc. Chem. Res.* **2008**, *41* (10), 1331–1342.

(220) Laufer, B.; Chatterjee, J.; Frank, A. O.; Kessler, H. *J. Pept. Sci.* **2009**, *15* (3), 141–146.

(221) Laufer, B.; Frank, A. O.; Chatterjee, J.; Neubauer, T.; Mas-Moruno, C.; Kummerlöwe, G.; Kessler, H. *Chem. - A Eur. J.* **2010**, *16* (18), 5385–5390.

(222) Wu, M.; Hancock, R. E. W. *J. Biol. Chem.* **1999**, *274* (1), 29–35.

(223) Sai, K. P.; Jagannadham, M. V.; Vairamani, M.; Raju, N. P.; Devi, A. S.; Nagaraj, R.; Sitaram, N. *J. Biol. Chem.* **2001**, *276* (4), 2701–2707.

(224) Salmon, A. L.; Cross, L. J. M.; Irvine, A. E.; Lappin, T. R. J.; Dathell, M.; Krausell, G.; Canning, P.; Thim, L.; Beyermann, M.; Rothemund, S.; Bienert, M.; Shaw, C. *J. Biol. Chem.* **2001**, *276* (13), 10145–10152.

(225) Shi, T.; Spain, S. M.; Rabenstein, D. L. *J. Am. Chem. Soc.* **2004**, *126* (3), 790–796.

(226) Shi, T.; Spain, S. M.; Rabenstein, D. L. *Angew. Chem. - Int. Ed.* **2006**, *45* (11), 1780–1783.

(227) Hörner, M.; Weber, W. *FEBS Lett.* **2012**, *586* (15), 2084–2096.

(228) Ambroggio, X. I.; Kuhlman, B. *Curr. Opin. Struct. Biol.* **2006**, *16* (4), 525–530.

(229) Mueller, C.; Grossmann, T. N. *Angew. Chem. - Int. Ed.* **2018**, *57* (52), 17079–17083.

(230) Zhou, X. Z.; Lu, P. J.; Wulf, G.; Lu, K. P. *Cell. Mol. Life Sci.* **1999**, *56* (9–10), 788–806.

(231) Schutkowski, M.; Bernhardt, A.; Zhou, X. Z.; Shen, M.; Reimer, U.; Rahfeld, J. U.; Lu, K. P.; Fischer, G. *Biochemistry* **1998**, *37* (16), 5566–5575.

(232) Signarvic, R. S.; DeGrado, W. F. *J. Mol. Biol.* **2003**, *334* (1), 1–12.

(233) Aemissegger, A.; Kräutler, V.; van Gunsteren, W. F.; Hilvert, D. *J. Am. Chem. Soc.* **2005**, *127* (9), 2929–2936.

(234) Nuti, F.; Gellini, C.; Larregola, M.; Squillantini, L.; Chelli, R.; Salvi, P. R.; Lequin, O.; Pietraperzia, G.; Papini, A. M. *Front. Chem.* **2019**, *7* (180), 1–9.

(235) Hoppmann, C.; Seedorff, S.; Richter, A.; Fabian, H.; Schmieder, P.; Rück-Braun, K.; Beyermann, M. *Angew. Chem. - Int. Ed.* **2009**, *48* (36), 6636–6639.

(236) Mart, R. J.; Allemann, R. K. *Chem. Commun.* **2016**, *52* (83), 12262–12277.

(237) Anderson, J. M.; Andersen, N. H. *Angew. Chem. - Int. Ed.* **2017**, *56* (25), 7074–7077.

(238) Awasthi, S. K.; Shankaramma, S. C.; Raghothama, S.; Balaram, P. *Biopolymers* **2001**, *58*, 465–476.

(239) Balamurugan, D.; Muraleedharan, K. M. *European J. Org. Chem.* **2015**, *2015* (24), 5321–5325.

(240) Dado, G. P.; Gellman, S. H. *J. Am. Chem. Soc.* **1993**, *115* (26), 12609–12610.

(241) Pandya, M. J.; Cerasoli, E.; Joseph, A.; Stoneman, R. G.; Waite, E.; Woolfson, D. N. *J. Am. Chem. Soc.* **2004**, *126* (51), 17016–17024.

(242) Jiang, L.; Burgess, K. *J. Am. Chem. Soc.* **2002**, *124* (31), 9028–9029.

(243) Bayse, C. A.; Pollard, D. B. *J. Pept. Sci.* **2019**, *25* (6), 16–22.

(244) Park, C.; Raines, R. T. *Protein Eng. Des. Sel.* **2001**, *14* (11), 939–942.

(245) Vallée-Bélisle, A.; Plaxco, K. W. *Curr. Opin. Struct. Biol.* **2010**, *20* (4), 518–526.

(246) He, L.; Navarro, A. E.; Shi, Z.; Kallenbach, N. R. *J. Am. Chem. Soc.* **2012**, *134* (3), 1571–1576.

(247) Cardamone, S.; Hughes, T. J.; Popelier, P. L. A. *Phys. Chem. Chem. Phys.* **2014**, *16* (22), 10367–10387.

(248) Klepeis, J. L.; Lindorff-Larsen, K.; Dror, R. O.; Shaw, D. E. *Curr. Opin. Struct. Biol.* **2009**, *19* (2), 120–127.

(249) Miller, S. C.; Scanlan, T. S. *J. Am. Chem. Soc.* **1997**, *119* (9), 2301–2302.

(250) Chatterjee, J.; Laufer, B.; Kessler, H. *Nat. Protoc.* **2012**, *7* (3), 432–444.

(251) Marraud, M.; Dupont, V.; Grand, V.; Zerkout, S.; Lecoq, A.; Boussard, G.; Vidal, J.; Collet, A.; Aubry, A. *Biopolymers* **1993**, *33* (7), 1135–1148.

(252) Sui, Q.; Rabenstein, D. L. *Pept. Sci.* **2018**, *e24088*, 1–8.

(253) Bacsa, B.; Horváti, K.; Bõsze, S.; Andreae, F.; Kappe, C. O. *J. Org. Chem.* **2008**, *73* (19), 7532–7542.

(254) Calce, E.; Vitale, R. M.; Scaloni, A.; Amodeo, P.; De Luca, S. *Amino Acids* **2015**, *47* (8), 1507–1515.

(255) Postma, T. M.; Albericio, F. *European J. Org. Chem.* **2014**, 3519–3530.

(256) Postma, T. M.; Albericio, F. *RSC Adv.* **2013**, *3* (34).

(257) Maruyama, K.; Nagasawa, H.; Suzuki, A. *Peptides* **1999**, *20* (7), 881–884.

(258) Hudáky, I.; Gáspári, Z.; Carugo, O.; Čemažar, M.; Pongor, S.; Perczel, A. *Proteins Struct. Funct. Bioinforma.* **2004**, *55* (1), 152–168.

(259) Petrotchenko, E. V.; Borchers, C. H. *Modern mass spectrometry-based structural proteomics*, 1st ed.; Elsevier Inc., 2014; Vol. 95.

(260) Chung, B. K. W.; Yudin, A. K. *Org. Biomol. Chem.* **2015**, *13* (33), 8768–8779.

(261) Otting, G.; Liepinsh, E.; Wüthrich, K. *Biochemistry* **1993**, *32* (14), 3571–3582.

(262) Crecente-Garcia, S.; Neckebroeck, A.; Clark, J. S.; Smith, B. O.; Thomson, A. R.

Org. Lett. **2020**, *22*, 4424–4428.

(263) Dirksen, A.; Yegneswaran, S.; Dawson, P. E. *Angew. Chem. - Int. Ed.* **2010**, *49* (11), 2023–2027.

(264) Senter, P. D. *Curr. Opin. Chem. Biol.* **2009**, *13* (3), 235–244.

(265) Nisal, R.; Jose, G.; Shanbhag, C.; Kalia, J. *Org. Biomol. Chem.* **2018**, *16* (23), 4304–4310.

(266) Takaoka, Y.; Tsutsumi, H.; Kasagi, N.; Nakata, E.; Hamachi, I. *J. Am. Chem. Soc.* **2006**, *128* (10), 3273–3280.

(267) Fang, G. M.; Wang, J. X.; Liu, L. *Angew. Chem. - Int. Ed.* **2012**, *51* (41), 10347–10350.

(268) Fang, G. M.; Li, Y. M.; Shen, F.; Huang, Y. C.; Li, J. Bin; Lin, Y.; Cui, H. K.; Liu, L. *Angew. Chem. - Int. Ed.* **2011**, *50* (33), 7645–7649.

(269) Wilkinson, A. A.; Jagu, E.; Ubych, K.; Coulthard, S.; Rushton, A. E.; Kennefick, J.; Su, Q.; Neely, R. K.; Fernandez-Trillo, P. *ACS Cent. Sci.* **2020**, *6* (4), 525–534.

(270) Aubry, A.; Bayeul, D.; Mangeot, J. -P; Vidal, J.; Sterin, S.; Collet, A.; Lecoq, A.; Marraud, M. *Biopolymers* **1991**, *31* (6), 793–801.

(271) Zerkout, S.; Dupont, V.; Aubry, A.; Vidal, J.; Collet, A.; Vicherat, A.; Marraud, M. *Int. J. Pept. Protein Res.* **1994**, *44*, 378–387.

(272) Poupart, J.; Hamdane, Y.; Lubell, W. D. *Can. J. Chem.* **2020**, *98* (6), 278–284.

(273) Thévenet, L.; Vanderesse, R.; Marraud, M.; Didierjean, C.; Aubry, A. *Tetrahedron Lett.* **2000**, *41* (14), 2361–2364.

(274) Yan, X. S.; Luo, H.; Zou, K. S.; Cao, J. L.; Li, Z.; Jiang, Y. B. *ACS Omega* **2018**, *3* (5), 4786–4790.

(275) Lascano, S.; Zhang, K.; Wehlauch, R.; Gademann, K.; Sakai, N.; Matile, S. *Chem. Sci.* **2016**, *7*, 4720–4724.

(276) Sreerama, N.; Woody, R. W. In *Methods in Enzymology*; 2004; Vol. 383, pp 318–351.

(277) Greenfield, N. J. *Nat. Protoc.* **2007**, *1* (6), 2527–2535.

(278) Reid, D. G. *Methods in Molecular Biology, volume 60: Protein NMR Techniques*; Humana Press: Totowa, New Jersey, 1997; Vol. 1.

(279) Williamson, M. P. *Methods in molecular biology, volume 17: Spectroscopic methods and analyses*; Jones, C., Mulloy, B., Thomas, A. H., Eds.; Humana Press: Totowa, New Jersey, 1993.

(280) Andersen, N. H.; Olsen, K. A.; Fesinmeyer, R. M.; Tan, X.; Hudson, F. M.; Eidenschink, L. A.; Farazi, S. R. *J. Am. Chem. Soc.* **2006**, *128* (18), 6101–6110.

(281) Hunter, C. A.; Packer, M. J.; Zonta, C. *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, *47* (1–2), 27–39.

(282) Wójcik, P.; Berlicki, Ł. *Bioorganic Med. Chem. Lett.* **2016**, *26* (3), 707–713.

(283) Mabonga, L.; Kappo, A. P. *Int. J. Pept. Res. Ther.* **2020**, *26* (1), 225–241.

(284) Jubb, H.; Blundell, T. L.; Ascher, D. B. *Prog. Biophys. Mol. Biol.* **2015**, *119* (1), 2–9.

(285) Wilson, A. J. *Chem. Soc. Rev.* **2009**, *38* (12), 3289–3300.

(286) Tsomaia, N. *Eur. J. Med. Chem.* **2015**, *94*, 459–470.

(287) Lee, A. C. L.; Harris, J. L.; Khanna, K. K.; Hong, J. H. *Int. J. Mol. Sci.* **2019**, *20* (10), 1–21.

(288) Rezaei Araghi, R.; Keating, A. E. *Curr. Opin. Struct. Biol.* **2016**, *39*, 27–38.

(289) Whitby, L. R.; Ando, Y.; Setola, V.; Vogt, P. K.; Roth, B. L.; Boger, D. L. *J. Am. Chem. Soc.* **2011**, *133* (26), 10184–10194.

(290) Philippe, G.; Huang, Y. H.; Cheneval, O.; Lawrence, N.; Zhang, Z.; Fairlie, D. P.; Craik, D. J.; de Araujo, A. D.; Henriques, S. T. *Biopolymers* **2016**, *106* (6), 853–863.

(291) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. *Science* **1996**, *274* (5289), 948–953.

(292) Bernal, F.; Tyler, A. F.; Korsmeyer, S. J.; Walensky, L. D.; Verdine, G. L. *J. Am. Chem. Soc.* **2007**, *129* (16), 2456–2457.

(293) Hara, T.; Durell, S. R.; Myers, M. C.; Appella, D. H. *J. Am. Chem. Soc.* **2006**, *128* (6), 1995–2004.

(294) Henchey, L. K.; Porter, J. R.; Ghosh, I.; Arora, P. S. *ChemBioChem* **2010**, *11* (15), 2104–2107.

(295) Kritzer, J. A.; Zutshi, R.; Cheah, M.; Ran, F. A.; Webman, R.; Wongjirad, T. M.; Schepartz, A. *ChemBioChem* **2006**, *7* (1), 29–31.

(296) Strizhak, A. V.; Babii, O.; Afonin, S.; Bakanovich, I.; Pantelejevs, T.; Xu, W.; Fowler, E.; Eapen, R.; Sharma, K.; Platonov, M. O.; Hurmach, V. V.; Itzhaki, L.; Hyvönen, M.; Ulrich, A. S.; Spring, D. R.; Komarov, I. V. *Org. Biomol. Chem.* **2020**, *18* (28), 5359–5369.

(297) Lau, Y. H.; Wu, Y.; Rossmann, M.; Tan, B. X.; De Andrade, P.; Tan, Y. S.; Verma, C.; McKenzie, G. J.; Venkitaraman, A. R.; Hyvönen, M.; Spring, D. R. *Angew. Chem. - Int. Ed.* **2015**, *54* (51), 15410–15413.

(298) Baek, S.; Kutchukian, P. S.; Verdine, G. L.; Huber, R.; Holak, T. A.; Lee, K. W.; Popowicz, G. M. *J. Am. Chem. Soc.* **2012**, *134* (1), 103–106.

(299) Fasan, R.; Dias, R. L. A.; Moehle, K.; Zerbe, O.; Obrecht, D.; Mittl, P. R. E.; Grütter, M. G.; Robinson, J. A. *ChemBioChem* **2006**, *7* (3), 515–526.

(300) Danelius, E.; Pettersson, M.; Bred, M.; Min, J.; Waddell, M. B.; Guy, R. K.; Grøtli, M.; Erdelyi, M. *Org. Biomol. Chem.* **2016**, *14* (44), 10386–10393.

(301) Kates, S. A.; Solé, N. A.; Johnson, C. R.; Hudson, D.; Barany, G.; Albericio, F. *Tetrahedron Lett.* **1993**, *34* (10), 1549–1552.

(302) Grieco, P.; Gitu, P. M.; Hruby, V. J. *J. Pept. Res.* **2001**, *57* (3), 250–256.

(303) Thakkar, A.; Trinh, T. B.; Pei, D. *ACS Comb. Sci.* **2013**, *15* (2), 120–129.

(304) Schasfoort, R. B. M. In *Handbook of Surface Plasmon Resonance*; Schasfoort, R. B.

M., Ed.; Royal Society of Chemistry, 2017; pp 1–26.

(305) SPR Assay Services for Drug Discovery | Reaction Biology https://www.reactionbiology.com/services/biophysical-assays/surface-plasmon-resonance-spr (accessed May 11, 2021).

(306) Douzi, B. In *Bacterial Protein Secretion Systems: Methods and Protocols*; Journet, L., Cascales, E., Eds.; Springer, 2017; Vol. 1615, pp 257–275.

(307) Chelushkin, P. S.; Leko, M. V.; Dorosh, M. Y.; Burov, S. V. *J. Pept. Sci.* **2017**, *23* (1), 13–15.

(308) Ulrich, S.; Boturyn, D.; Marra, A.; Renaudet, O.; Dumy, P. *Chem. - A Eur. J.* **2014**, *20* (1), 34–41.

(309) Sohma, Y.; Kent, S. B. H. *J. Am. Chem. Soc.* **2009**, *131* (44), 16313–16318.

(310) Dirksen, A.; Dawson, P. E. *Bioconjug. Chem.* **2008**, *19* (12), 2543–2548.

(311) Kool, E. T.; Crisalli, P.; Chan, K. M. *Org. Lett.* **2014**, *16* (5), 1454–1457.

(312) Larsen, D.; Pittelkow, M.; Karmakar, S.; Kool, E. T. *Org. Lett.* **2015**, *17* (2), 274–277.

(313) Kool, E. T.; Park, D. H.; Crisalli, P. *J. Am. Chem. Soc.* **2013**, *135* (47), 17663–17666.

(314) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J. L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, *106* (9), 3652–3711.

(315) Agten, S. M.; Suylen, D.; Ippel, H.; Kokozidou, M.; Tans, G.; Van De Vijver, P.; Koenen, R. R.; Hackeng, T. M. *ChemBioChem* **2013**, *14* (18), 2431–2434.

(316) Guthrie, Q. A. E.; Young, H. A.; Proulx, C. *Chem. Sci.* **2019**, *10* (41), 9506–9512.

(317) Roberts, K. D.; Lambert, J. N.; Ede, N. J.; Bray, A. M. *J. Pept. Sci.* **2004**, *10* (11), 659–665.

(318) Lamping, M.; Grell, Y.; Geyer, A. *J. Pept. Sci.* **2016**, *22* (4), 228–235.

(319) Duflocq, S.; Zhou, J.; Huguenot, F.; Vidal, M.; Liu, W. Q. *RSC Adv.* **2020**, *10* (30), 17681–17685.

(320) Bi, X.; Pasunooti, K. K.; Lescar, J.; Liu, C. F. *Bioconjug. Chem.* **2017**, *28* (2), 325–329.

(321) Spetzler, J. C.; Hoeg-Jensen, T. *J. Pept. Sci.* **2001**, *7* (10), 537–551.

(322) Rashidian, M.; Mahmoodi, M. M.; Shah, R.; Dozier, J. K.; Wagner, C. R.; Distefano, M. D. *Bioconjug. Chem.* **2013**, *24* (3), 333–342.

(323) Wendeler, M.; Grinberg, L.; Wang, X.; Dawson, P. E.; Baca, M. *Bioconjug. Chem.* **2014**, *25* (1), 93–101.

(324) Wang, S.; Gurav, D.; Oommen, O. P.; Varghese, O. P. *Chem. - A Eur. J.* **2015**, *21* (15), 5980–5985.

(325) Hogg, J. L.; Jencks, D. A.; Jencks, W. P. *J. Am. Chem. Soc.* **1977**, *99* (14), 4772–4778.

(326) Fischer, H.; DeCandis, F. X.; Ogden, S. D.; Jencks, W. P. *J. Am. Chem. Soc.* **1980**,

102 (4), 1340–1347.

(327) Pallin, T. D.; Tam, J. P. *J. Chem. Soc. Chem. Commun.* **1995**, *8* (19), 2021–2022.

(328) Haney, C. M.; Horne, W. S. *J. Pept. Sci.* **2014**, *20* (2), 108–114.

(329) Agten, S. M.; Dawson, P. E.; Hackeng, T. M. *J. Pept. Sci.* **2016**, *22* (5), 271–279.

(330) Lelièvre, D.; Buré, C.; Laot, F.; Delmas, A. *Tetrahedron Lett.* **2001**, *42* (2), 235–238.

(331) Avan, I.; Dennis Hall, C.; Katritzky, A. R. *Chem. Soc. Rev.* **2014**, *43* (10), 3575–3594.

(332) Degruyter, J. N.; Malins, L. R.; Wimmer, L.; Clay, K. J.; Lopez-Ogalla, J.; Qin, T.; Cornella, J.; Liu, Z.; Che, G.; Bao, D.; Stevens, J. M.; Qiao, J. X.; Allen, M. P.; Poss, M. A.; Baran, P. S. *Org. Lett.* **2017**, *19* (22), 6196–6199.

(333) Yang, Y. *Side reactions in peptide synthesis*; Academic Press, 2016.

(334) Yun, F.; Cheng, C.; Ullah, S.; Yuan, Q. *Eur. J. Med. Chem.* **2020**, *198*, 112322.

(335) Soto, S.; Vaz, E.; Dell'Aversana, C.; Álvarez, R.; Altucci, L.; De Lera, Á. R. *Org. Biomol. Chem.* **2012**, *10* (10), 2101–2112.

(336) Peacock, L. R.; Chapman, R. S. L.; Sedgwick, A. C.; Mahon, M. F.; Amans, D.; Bull, S. D. *Org. Lett.* **2015**, *17* (4), 994–997.

(337) Clayden, J.; Pickworth, M.; Jones, L. H. *Chem. Commun.* **2009**, No. 5, 547–549.

(338) Kudoh, T.; Park, C. S.; Lefurgy, S. T.; Sun, M.; Michels, T.; Leyh, T. S.; Silverman, R. B. *Bioorganic Med. Chem.* **2010**, *18* (3), 1124–1134.

(339) Pitts, M. R.; Harrison, J. R.; Moody, C. J. *J. Chem. Soc. Perkin Trans. 1* **2001**, No. 9, 955–977.

(340) Orlandi, M.; Tosi, F.; Bonsignore, M.; Benaglia, M. *Org. Lett.* **2015**, *17* (16), 3941–3943.

(341) Kumar, N. N. B.; Kutateladze, A. G. *Org. Lett.* **2016**, *18* (3), 460–463.

(342) Jiao, L.; Qiu, Q.; Liu, B.; Zhao, T.; Huang, W.; Qian, H. *Bioorganic Med. Chem.* **2014**, *22* (24), 6857–6866.

(343) Shi, L.; Hu, R.; Wei, Y.; Liang, Y.; Yang, Z.; Ke, S. *Eur. J. Med. Chem.* **2012**, *54*, 549–556.

(344) Dignam, K. J.; Hegarty, A. F. *J. Chem. Soc. Perkin Trans. 2* **1979**, 1437–1443.

(345) Johnson, J. E.; Morales, N. M.; Gorczyca, A. M.; Dolliver, D. D.; McAllister, M. A. *J. Org. Chem.* **2001**, *66* (24), 7979–7985.

(346) Yang, J.; Zhou, S.; Ji, L.; Zhang, C.; Yu, S.; Li, Z.; Meng, X. *Bioorg. Med. Chem. Lett.* **2014**, *24* (21), 5055–5058.

(347) Cistrone, P. A.; Dirksen, A.; Ingale, S.; Dawson, P. E. *Aust. J. Chem.* **2020**, *73* (4), 377–379.

(348) Shen, L.; Cao, N.; Tong, L.; Zhang, X.; Wu, G.; Jiao, T.; Yin, Q.; Zhu, J.; Pan, Y.; Li, H. *Angew. Chem. - Int. Ed.* **2018**, *57* (50), 16486–16490.

(349) Hegedüs, Z.; Hóbor, F.; Shoemark, D. K.; Celis, S.; Lian, L. Y.; Trinh, C. H.;

Sessions, R. B.; Edwards, T. A.; Wilson, A. J. *Chem. Sci.* **2021**, *12* (6), 2286–2293.

(350) Touw, W. G.; Baakman, C.; Black, J.; Te Beek, T. A. H.; Krieger, E.; Joosten, R. P.; Vriend, G. *Nucleic Acids Res.* **2015**, *43* (D1), D364–D368.

(351) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeerschd, T.; Zurek, E.; Hutchison, G. R. *J. Cheminform.* **2012**, *4* (17), 1–17.

(352) Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4* (3), 435–447.

(353) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. *Comput. Phys. Commun.* **2014**, *185* (2), 604–613.

(354) Laio, A.; Gervasio, F. L. *Reports Prog. Phys.* **2008**, *71* (12).

(355) D.A. Case, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T. A. D.; R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, Y. H.; S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. J.; Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R.; Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. SalomonFerrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D. M. Y. and P. A. K. *AMBER 2018*. 2018.

(356) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. *PLOS Comput. Biol.* **2017**, *13* (7), e1005659.

(357) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. *J. Comput. Chem.* **2011**, *32* (10), 2319–2327.

(358) Dou, H.; Buetow, L.; Hock, A.; Sibbet, G. J.; Vousden, K. H.; Huang, D. T. *Nat. Struct. Mol. Biol.* **2012**, *19* (2), 184–192.

(359) Skinner, S. P.; Fogh, R. H.; Boucher, W.; Ragan, T. J.; Mureddu, L. G.; Vuister, G. W. *J. Biomol. NMR* **2016**, *66* (2), 111–124.

(360) Annis, I.; Hargittai, B.; Barany, G. *Methods Enzymol.* **1997**, *289*, 198–221.

(361) Beard, R.; Stucki, A.; Schmitt, M.; Py, G.; Grundschober, C.; Gee, A. D.; Tate, E. W. *Bioorganic Med. Chem.* **2018**, *26* (11), 3039–3045.

(362) Postma, T. M.; Albericio, F. *RSC Adv.* **2013**, *3* (34), 14277–14280.

(363) Tam, J. P.; Wu, C. R.; Liu, W.; Zhang, J. W. *J. Am. Chem. Soc.* **1991**, *113* (17), 6657–6662.

(364) Zhu, S.; Xu, B.; Wang, G.; Zhou, Q. *J. Am. Chem. Soc* **2012**, *134* (1), 436–442.

(365) Kelly, S. M.; Jess, T. J.; Price, N. C. *Biochim. Biophys. Acta - Proteins Proteomics* **2005**, *1751* (2), 119–139.

(366) Mukhina, O. A.; Bhuvan Kumar, N. N.; Arisco, T. M.; Valiulin, R. A.; Metzel, G. A.; Kutateladze, A. G. *Angew. Chem. - Int. Ed.* **2011**, *50* (40), 9423–9428.

(367) Sun, R.; Li, Y.; Lü, M.; Xiong, L.; Wang, Q. *Bioorganic Med. Chem. Lett.* **2010**, *20* (15), 4693–4699.

(368) Günther, M.; Lategahn, J.; Juchum, M.; Döring, E.; Keul, M.; Engel, J.; Tumbrink,

H. L.; Rauh, D.; Laufer, S. *J. Med. Chem.* **2017**, *60* (13), 5613–5637.